

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 09:04:55 ; Search time 5325 Seconds  
(without alignments)  
10670.933 Million cell updates/sec

Title: US-10-081-739A-1

Perfect score: 1311

Sequence: 1 atggccaagtactccgagct.....gctactcgggggtgggctga 1311

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sv.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rtd.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sv.\*

39: em\_higo\_hum.\*

40: em\_higo\_mus.\*

41: em\_higo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1311	100.0	1311	12	AF504065	AF504065 Synthetic
2	1102.6	84.1	1299	13	AF504064	AF504064 Uncultured
3	1085	82.8	1386	1	AF504062	AF504062 Thermococ
4	1040.8	79.4	1386	1	AF504063	AF504063 Thermococ
5	1003.4	76.5	2705	1	AF068255	AF068255 Thermococ
6	991.2	75.6	2179	1	D83793	D83793 Pyrococcus
7	989.6	75.5	2179	6	E13334	E13334 GDNA encodi
8	963.6	73.5	2372	1	AF017454	AF017454 Thermococ
9	792.8	60.5	1480	1	AF177906	AF177906 Thermococ
10	792.8	60.5	1561	1	AF240464	AF240464 Pyrococcus
11	792.8	60.5	1740	1	PF096622	U96822 Pyrococcus
12	792.8	60.5	2134	6	E09372	E09372 Heat-resist
13	792.8	60.5	2627	1	AF001268	AF001268 Pyrococcus
14	792.8	60.5	3519	6	BD012151	BD012151 Polypepti
15	792.8	60.5	10643	1	AE010170	AE010170 Pyrococcus
16	768.2	58.6	1206	6	E14254	E14254 Thermococ
17	146.4	11.2	221	1	AF068256	AF068256 Thermococ
18	87	6.6	12635	1	AE009342	AE009342 Agrobacte
19	87	6.6	13095	1	AE008292	AE008292 Agrobacte
20	85.2	6.5	245050	1	AL627272	AL627272 Salmonell
21	85.2	6.5	300247	1	AE016837	AE016837 Salmonell
22	83.6	6.4	1700	1	STVFTAA	L01643 S.typhimuri
23	83.6	6.4	24156	1	AE008787	AE008787 Salmonell
24	82	6.3	334520	1	AP003588	AP003588 Nostoc sp
25	79.6	6.1	1458	6	AR175875	AR175875 Sequence
26	79.6	6.1	1458	6	BD249239	BD249239 Polypepti
27	79.6	6.1	1458	6	AR401551	AR401551 Sequence
28	79.6	6.1	1458	6	AX037042	AX037042 Sequence
29	78.4	6.0	1539	1	AF438149	AF438149 Bacillus
30	76.8	5.9	1449	6	A23402	A23402 B.lichenifo
31	76.8	5.9	1449	6	A27772	A27772 Alpha amyla
32	76.8	5.9	1452	6	AX705190	AX705190 Sequence
33	76.8	5.9	1536	6	E01158	E01158 DNA sequenc
34	76.8	5.9	1539	6	E12201	E12201 DNA encodin
35	76.8	5.9	1777	6	A17930	A17930 Alpha amyla
36	76.8	5.9	1777	6	I24553	I24553 Sequence 33
37	76.8	5.9	1777	6	I83686	I83686 Sequence 16
38	76.8	5.9	1948	1	BACAMYABL	M38570 B.lichenifo
39	76.8	5.9	1948	1	BLAAMYLG	X03236 Bacillus 11
40	76.8	5.9	1968	1	BACAMYS	M13256 B.lichenifo
41	76.8	5.9	1968	6	A91241	A91241 Sequence 33
42	76.8	5.9	1968	6	AR000567	AR000567 Sequence
43	76.8	5.9	1968	6	AR012365	AR012365 Sequence
44	76.8	5.9	1968	6	AR049549	AR049549 Sequence
45	76.8	5.9	1968	6	AR065754	AR065754 Sequence

# ALIGNMENTS

RESULT 1	AF504065	AF504065	1311 bp	DNA	linear	SYN 15-JUL-2002
LOCUS	AF504065	Synthetic construct alpha-amylase precursor (BD5088) gene, complete cds.				
DEFINITION	AF504065	AF504065.1	GI:21327000			
ACCESSION	AF504065					
VERSION	AF504065					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						

Pred. No. is the number of results predicted by chance to have a

JOURNAL  
PUBLISHED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES

AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE  
J. Biol. Chem. 277 (29), 26501-26507 (2002)  
11994309  
2 (bases 1 to 1311)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabell,M., Lam,D.,  
Macomber,J., Short,J.M., Robertson,D.E. and Miller,C.  
Direct Submission  
Submitted (18-APR-2002) Bioinformatics, Diversa Corporation, 4955  
Directors Place, San Diego, CA 92121, USA

LOCATION/Qualifiers

1..1311

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

1..1311

/gene="BD5088"

1..1311

/gene="BD5088"

/codon\_start=1

/transl\_table=11

/product="alpha-amylase precursor"

/protein\_id="AAM48115.1"

/db\_xref="GI:21327001"

/translation="MAKYSLEKGGVIMQAFYVDVPSGGIWDITROKIPWYDAGIS  
AIVPASKGMGAYSMGYDPDFDLGEYDQKGVETRFPSKQELVKNMINTAHAYGM  
KVIADIVINHRAGDLEWNPVNDYTWDFSVAGKYNTANYLDHFNELHAGDSGTF  
GGYPDICHKSNQYWLWASQSYAALRSIGIDAWRFYKGYAPWVYKDWLNMWGG  
WAVEGYDNTDAVNLWYSSGAKVDFALYIMDEAFDNKNI PALVSALQNGQTWVS  
RDPFKAVTVNHDITDIWKYPAVAFILTYEGQPTI FYRDEEWMNKDKLNLIIWH  
ENLAGGSQTDIVYDNDLIFVRNGYDKPLGLTYINLGSKAGRYVVPKPGACIHE  
YTNLGGWVDKYVYSSGWIYLEAPVDPANGQYGVSWSYCVGV"

ORIGIN

Query Match 100.0%; Score 1311; DB 12; Length 1311;

Best Local Similarity 100.0%; Pred. No. 1e-285;

Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCCAAGTACTCCGAGCTGGAAAGGCGGGGTCTATATGACGGCTTCTACTGGGAC 60
DB 1 ATGCCAAGTACTCCGAGCTGGAAAGGCGGGGTCTATATGACGGCTTCTACTGGGAC 60
QY 61 GTCCCTTCAGAGGAATATGTTGGGACACAATAACGGCAGAAGATACCGGAGTGGTACGAT 120
DB 61 GTCCCTTCAGAGGAATATGTTGGGACACAATAACGGCAGAAGATACCGGAGTGGTACGAT 120
QY 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGCGCCTATTCCG 180
DB 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGCGCCTATTCCG 180
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGGAACGGTA 240
DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGGAACGGTA 240
QY 241 GAGACGGCTTTGGTCCAGAGGAGCTCGTGAACATGATATAACACCCGCCACGCTAT 300
DB 241 GAGACGGCTTTGGTCCAGAGGAGCTCGTGAACATGATATAACACCCGCCACGCTAT 300
QY 301 GGCATGAAGGTAATAGCCGATATAGTCATCAACCCCGCGCGGTGACCTGGAGTGG 360
DB 301 GGCATGAAGGTAATAGCCGATATAGTCATCAACCCCGCGCGGTGACCTGGAGTGG 360
QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGCGTGGGTAATAC 420
DB 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGCGTGGGTAATAC 420
QY 421 ACGGCCAATCTACCTGACCTCCACCGGACGAGCTCCATGCGGGCGATTCCGGACATTT 480
DB 421 ACGGCCAATCTACCTGACCTCCACCGGACGAGCTCCATGCGGGCGATTCCGGACATTT 480
QY 481 GGAGGGTATCCCGACATATGCCACGACAAAGAGCTGGGACCACTACTGGCTCTGGGCGCAGC 540
DB 481 GGAGGGTATCCCGACATATGCCACGACAAAGAGCTGGGACCACTACTGGCTCTGGGCGCAGC 540

```

```

QY 541 CAGGAGAGCTACGGCGCATATCTCAGGAGCATCGGCATCGATCGCTGCGCTTCGACTAC 600
DB 541 CAGGAGAGCTACGGCGCATATCTCAGGAGCATCGGCATCGATCGCTGCGCTTCGACTAC 600
QY 601 GTCAAGGGCTATGCTCCTCGGTGCTCAAGGACTGGGCTGAACCTGGTGGGAGGCTGGCG 660
DB 601 GTCAAGGGCTATGCTCCTCGGTGCTCAAGGACTGGGCTGAACCTGGTGGGAGGCTGGCG 660
QY 661 GTTGAGAGTACTCGGACACCAACGCTGACGCTGTTCTCAACTGGGCATACTCGAGCGGT 720
DB 661 GTTGAGAGTACTCGGACACCAACGCTGACGCTGTTCTCAACTGGGCATACTCGAGCGGT 720
QY 721 GCCAAGGCTTTGACTTCGGCCCTCTACTACAAGATGATGAGCGCTTTTGACACAAAAAC 780
DB 721 GCCAAGGCTTTGACTTCGGCCCTCTACTACAAGATGATGAGCGCTTTTGACACAAAAAC 780
QY 781 ATTCGAGCGCTGCTCTGCGCTTCAGAACGGCGAGACTGTTCTCCCGGACCGCTTC 840
DB 781 ATTCGAGCGCTGCTCTGCGCTTCAGAACGGCGAGACTGTTCTCCCGGACCGCTTC 840
QY 841 AAGCCCGTAACCTTTGTAGCAACACACGACATATATCTGGAACAAGTATCCAGCC 900
DB 841 AAGCCCGTAACCTTTGTAGCAACACACGACATATATCTGGAACAAGTATCCAGCC 900
QY 901 TACCGGTTTCATCTCCTACCTAGGAGGCGCGGCACATATCTACCGGACTACGAGGAG 960
DB 901 TACCGGTTTCATCTCCTACCTAGGAGGCGCGGCACATATCTACCGGACTACGAGGAG 960
QY 961 TGGCTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATAGAACCTTCGCGGAGGA 1020
DB 961 TGGCTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATAGAACCTTCGCGGAGGA 1020
QY 1021 AGCACCGACATAGTCTACTAGGATACGATGAATCATCTCTGTCAGGAACGCTACGGG 1080
DB 1021 AGCACCGACATAGTCTACTAGGATACGATGAATCATCTCTGTCAGGAACGCTACGGG 1080
QY 1081 GACAAGCGGGGCTTAACTACATCAACTAGGCTCGAGCAAGCGCGGAAGGTGGGTT 1140
DB 1081 GACAAGCGGGGCTTAACTACATCAACTAGGCTCGAGCAAGCGCGGAAGGTGGGTT 1140
QY 1141 TATGTCCGAAGTTCGCGGCGGTGTCATCCAGGATATCTGGTAACCTTCGAGGCTGG 1200
DB 1141 TATGTCCGAAGTTCGCGGCGGTGTCATCCAGGATATCTGGTAACCTTCGAGGCTGG 1200
QY 1201 GTACAGAGTACGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
DB 1201 GTACAGAGTACGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
QY 1261 GCCAACGGGAGTATGCTACTCCGTGTGAGCTACTTCGCGGGGTGGGCTGA 1311
DB 1261 GCCAACGGGAGTATGCTACTCCGTGTGAGCTACTTCGCGGGGTGGGCTGA 1311

```

RESULT 2  
AF504064

LOCUS

DEFINITION

AF504064.1

GI:21326998

ACCESSION

AF504064

VERSION

AF504064.1

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

Uncultured organism alpha-amylase precursor (BD5063) gene, complete cds.  
1 (bases 1 to 1299)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabell,M., Lam,D.,  
Macomber,J., Short,J.M., Robertson,D.E. and Miller,C.  
A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
J. Biol. Chem. 277 (29), 26501-26507 (2002)  
11994309  
2 (bases 1 to 1299)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabell,M., Lam,D.,

Macomber, J., Short, J.M., Robertson, D.E. and Miller, C.  
 Direct Submission  
 Submitted (18-APR-2002) Bioinformatics, Diversa Corporation, 4955  
 Directors Place, San Diego, CA 92121, USA  
 Location/Qualifiers  
 source  
 1. .1299  
 /organism="uncultured organism"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:155900"  
 /note="cloned directly from environmental biomass DNA  
 isolated from a deep sea enrichment"  
 1. .1299  
 /gene="BD5063"  
 1. .1299  
 /gene="BD5063"  
 /codon\_start=1  
 /product="alpha-amylase precursor"  
 /protein\_id="AA048114.1"  
 /db\_xref="GI:21326999"  
 /translation="MALEEGLIMQAFYDFWPMGGIWDNIAQKIPDNWASAGISAIWI  
 FPASKGMSGYMDYDFDLEGYKQETRFSGKQELINIMNTAHAYGMKVI  
 DIVINHRAGDLEWPFVNDYTWDFSKVSGRYTANYLDFHPNELHAGDSGTFGGYP  
 DICDKWDQVLMWASOESVAAVLRISIGIDAMRFDYVKYAPVWVMDLWNGWMAVG  
 EYMTNDAVLNWAYSSGKVFDFALYKDEAFDNNIIPALVDALRYGOTVVSROPF  
 KAVTFVANHDTDIINWYPAVAPILLTYEGOPTIFRYDERELNKLKILWIDHDLA  
 GGSIDIYYENDELIFVRNGYSGKPLITINLASKAGRWVYVPFAGSCIHYETGN  
 LGGWDRWDSGWWILEAPAHDPANGQIGISWVSCGVG"  
 ORIGIN  
 Query Match 84.1%; Score 1102.6; DB 13; Length 1299;  
 Best Local Similarity 90.8%; Pred. No. 1.3e-238;  
 Matches 1174; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
 QY 19 CTGAAAGGCGGGGTCATATGCGAGCGTTCCTACTGGGACGTGCCTTCAGGAGGAATA 78  
 DB 7 CTGAAAGGCGGGGTCATATGCGAGCGTTCCTACTGGGACGTGCCTTCAGGAGGAATA 66  
 QY 79 TGGTGGGACAAATACGCGAGAGATACCGAGTGTACGATCCGGAATCTCCGCAATA 138  
 DB 67 TGGTGGGACAAATACGCGAGAGATACCGAGTGTACGATCCGGAATCTCCGCAATA 126  
 QY 139 TGGATTCCTCCGCGGAGCAAGGCGATGGCGCGCTTATTCGATGGGTACGACCCCTAC 198  
 DB 127 TGGATTCCTCCGCGGAGCAAGGCGATGAGCGCGCTTATTCGATGGGTACGACCCCTAC 186  
 QY 199 GACTTCCTTACCTCGGTAGTACGACGAGGAGGAGGTAGAGCGCGTTCGCTCC 258  
 DB 187 GATTATTTTGACCTCGGTAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 246  
 QY 259 AAGCAGGAGCTCGTGAACATGATAAACACCGCCACGCTATGGCATGAAGGTAATAGCC 318  
 DB 247 AAGCAGGAGCTCATAAACATGATAAACACCGCCACGCTATGGCATGAAGGTAATAGCC 306  
 QY 319 GATATAGTCAACACACGCGCGCGGTGACCTGGAGTGGAACCCCTTCGTGAACGAC 378  
 DB 307 GATATAGTCAACACACGCGCGCGGTGACCTGGAGTGGAACCCCTTCGTGAACGAC 366  
 QY 379 TATACCTGGACGACTTCTCAAAGTTCGCTGGGTAAATACAGGCCAATCTACTCGAC 438  
 DB 367 TATACCTGGACGACTTCTCAAAGTTCGCTGGGTAAATACAGGCCAATCTACTCGAC 426  
 QY 439 TTCCACCGGACGAGCTCCATCGGGCGATTCCGGAACATTTGGAGGCTATCCGACATA 498  
 DB 427 TTCCACCGGACGAGCTCCATCGGGCGATTCCGGAACATTTGGAGGCTATCCGACATA 486  
 QY 499 TCCACGACAGAGCTGGGACAGTACTGGCTCTGGCCAGCCAGGAGAGCTACGGGCA 558  
 DB 487 TCCACGACAGAGCTGGGACAGTACTGGCTCTGGCCAGCCAGGAGAGCTACGGGCA 546  
 QY 559 TATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAGCTCAAGGGCTATGCTCCC 618  
 DB 547 TATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAGCTCAAGGGCTATGCTCCC 606

QY 619 TGGTCTGTCAGGAGCTGGCTGAACCTGGTGGGAGGCTGGCGGTTGGAGAGTACTGGGAC 678  
 DB 607 TGGTCTGTCAGGAGCTGGCTGAACCTGGTGGGAGGCTGGCGGTTGGAGAGTACTGGGAC 666  
 QY 679 ACCAAGCTCGAGCGTGTCTCAACTGGGCATATCTCGAGCGGTGCCAAGTCTTTGACTTC 738  
 DB 667 ACCAAGCTCGAGCGTGTCTCAACTGGGCATATCTCGAGCGGTGCCAAGTCTTTGACTTC 726  
 QY 739 GCCCTCTACTACAAGATGGATGAGCGCTTTGACAAACAAACATTTCCAGCGGTCTCTCT 798  
 DB 727 GCCCTCTACTACAAGATGGATGAGCGCTTTGACAAACAAACATTTCCCGCCCTGGTGGAC 786  
 QY 799 GCCCTTCAGAACGGCGCAGACTGTCTCTCCCGGACCGCTTCAAGACCGTAACCTTTGTA 858  
 DB 787 GCCCTTCAGAACGGCGCAGACTGTCTCTCCCGGACCGCTTCAAGACCGTGTGACGTTGTA 846  
 QY 859 GCAAAACACGACACCGCATATAATCTGGAACAAAGTATCCAGCCTACCGGTTCTCTCACC 918  
 DB 847 GCAAAACACGATACCGCATATAATCTGGAACAAAGTATCCAGCCTACCGGTTCTCTCACC 906  
 QY 919 TAGGAGGGCGACCGCAATATTTACCGCGACTAGGAGGTGGCTCAACAGGATAAG 978  
 DB 907 TAGGAGGGCGACCGCAATATTTACCGCGACTAGGAGGTGGCTCAACAGGATAAG 966  
 QY 979 CTCAAGAACTCATCTGGATACATGAGAACCTTCGCGGAGGAGGAGGAGGAGGAGGAGGAG 1038  
 DB 967 CTCAAGAACTCATCTGGATACATGAGAACCTTCGCGGAGGAGGAGGAGGAGGAGGAGGAG 1026  
 QY 1039 TAGGATAACGATGAATCATCTTCTGTCAGGAACCGCTACGGGGAACAAGCCGGGCTTATA 1098  
 DB 1027 TAGGATAACGATGAATCATCTTCTGTCAGGAACCGCTACGGGGAACAAGCCGGGACTGATA 1086  
 QY 1099 ACTTACATCAACTAGGCTCGAGCAAGCGGGAAGTGGGTTTATGTGCGGAAGTTCGGG 1158  
 DB 1087 ACATACATCAACTCGCTCAAGCAAGCGGAGGTGGTTTACGTTCCGAGTTCGCA 1146  
 QY 1159 GGGCGTGCATCCAGGATATCTGTTAACTCCGAGGCTGGGTAGACAAGTACGCTCTAC 1218  
 DB 1147 GGGCTGTGCATACAGGATACACCGGCAATCTCGCGGCTGGGTGACAAAGTGGGTGGAC 1206  
 QY 1219 TCAAGCGGTGGGTCTATCTGAAAGTCCAGCTTACGACCTTCCGCAAGCGGAGTATGGC 1278  
 DB 1207 TCAAGCGGTGGGTCTATCTGAAAGTCCAGCTTACGACCTTCCGCAAGCGGAGTATGGC 1266  
 QY 1279 TACTCCGTGTGAGCTACTGCGGGGTGGGTGA 1311  
 DB 1267 TACTCCGTGTGAGCTACTGCGGGGTGGGTGA 1299  
 RESULT 3  
 AF504062  
 LOCUS  
 DEFINITION  
 Thermococcus sp. GU5L5 alpha-amylase precursor (BD5064) gene,  
 complete cds.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Thermococcus sp. GU5L5  
 Thermococcus sp. GU5L5  
 Archaea; Euryarchaeota; Thermococci; Thermococcales;  
 Thermococcaceae; Thermococcus.  
 REFERENCE  
 1 (bases 1 to 1386)  
 Richardson, T.H., Tan, X., Frey, G., Callen, W., Cabell, M., Lam, D.,  
 Macomber, J., Short, J.M., Robertson, D.E. and Miller, C.  
 A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
 AND OPTIMIZATION OF A LOW PH, THERMOSTABLE alpha -AMYLASE  
 J. Biol. Chem. 277 (29), 26501-26507 (2002)  
 11994309  
 2 (bases 1 to 1386)  
 Richardson, T.H., Tan, X., Frey, G., Callen, W., Cabell, M., Lam, D.,  
 Macomber, J., Short, J.M., Robertson, D.E. and Miller, C.  
 Direct Submission  
 Submitted (18-APR-2002) Bioinformatics, Diversa Corporation, 4955

Directors Place, San Diego, CA 92121, USA

# FEATURES

Location/Qualifiers  
1..1386  
/organism="Thermococcus sp. GU515"  
/mol\_type="genomic DNA"  
/strain="GU515"  
/db\_xref="taxon:195100"  
1..1386  
/gene="BD5064"  
1..1386  
/gene="BD5064"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-amylase precursor"  
/protein\_id="AAM48112.1"  
/db\_xref="GI:21326995"  
/translation="MKPAKLLVFLVLSILAGLYAQPAGAAKYLEBEGGVIMQFYW  
DVPSSGIWDDTIRQKIPYDAGISAIWIPPKSGMGAYSMGYDPYDFDLGEYDQK  
GTVEYFGSKQLVNMINTAHAYGIVKVIADIVINHRAGDLEWPFVNDYTWTFPSKV  
ASGKYTANYLDFHPNEVKDCDEGTGGPDIAHEKSDQYWLWASNEVYAAVLRISGV  
DAWRPDYKYGAVVVKDMLDMWGWAGEYWDVNDALLNWAYSSDAKVDFFPLYYK  
MDAAFDNKIPALVEALKNGTVVSDDPEKAVTFVANHDTDLIWNKYPAFAPILTYEG  
QPTIFRDYEELNDRLNKLIWIHDHLAGGSTDIVYDNDLELIFVRNGYDGKPLGIT  
YINLSSKAGRWYVPKFAAGACIHEYTNLGNLWVDKWDSSGWVYLEAPADHPANGYY  
GYSVWSYCGVG"  
1..78  
/gene="BD5064"

sig\_peptide

## ORIGIN

Query Match 82.8%; Score 1085; DB 1; Length 1386;  
Best Local Similarity 89.3%; Pred. No. 1.2e-234;  
Matches 1169; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

3 GGCCTAGTACTCGAGCTGGAAGGGCGGCTCATATGACGGGCTTCTACTGGAGCT 62  
78 GGCCTAGTACTCGAGCTGGAAGGGCGGCTCATATGACGGGCTTCTACTGGAGCT 137  
63 GCCTTCAGGAGGAATATGTTGGACACAATACGGCAGAGATACCGGAGTGGTACGATGC 122  
138 GCCTTCAGGAGGAATATGTTGGACACAATACGGCAGAGATACCGGAGTGGTACGATGC 197  
123 CGGAATCTCGCAATATGATATCCCGCGGAGCAGGCGATGGCGGCGCTATTCCGAT 182  
198 CGGAATCTCGCAATATGATATCCCGCGGAGCAGGCGATGGCGGCGCTATTCCGAT 257  
183 GGGCTACGACCCCTACGACTCTTTTGAACCTCGGTGAGTACGACCAAGAGGAAACGGTAGA 242  
258 GGGCTACGACCCCTACGACTCTTTTGAACCTCGGTGAGTACGACCAAGAGGAAACGGTAGA 317  
243 GAGCGCTTGGCTCAAGCAGGAGCTGCTGAACATGATAAACAACCGCCGACGCTATGG 302  
318 GAGCGCTTGGCTCAAGCAGGAGCTGCTGAACATGATAAACAACCGCCGACGCTATGG 377  
303 CATGAAGGTAATAGCGATATAGTATCAACCAACCGCGCGGCTGACCTGGAGTGGAA 362  
378 CATGAAGGTCATCGCAGACATAGTATCAACCAACCGCGCGGAGGAGACCTTGGTGA 437  
363 CCGCTTCGTGACGACTATACCTGACCGGACTTCTCAAGGTGCGGTGGGTAATAACAC 422  
438 CCGCTTCGTGACGACTATACCTGACCGGACTTCTCAAGGTGCGGTGGGTAATAACAC 497  
423 GGCCTACCTACCTCGACTTCCACCCGAAACGAGCTTCAATCGGCGGATTCGGAAACATTTGG 482  
498 GGCCTACCTACCTCGACTTCCACCCGAAACGAGCTTCAATCGGCGGATTCGGAAACATTTGG 557  
483 AGGCTATCCCGACATATGCGACCAAGAGCTGGACCGAGTACTGGCTCTGGCGGAGCA 542  
558 AGGCTATCCCGACATATGCGACCAAGAGCTGGACCGAGTACTGGCTCTGGCGGAGCA 617  
543 GAGAGAGTACGCGGATATCTCAGGAGGATCGGATCGATCGCTGGCGCTTCGACTAGCT 602  
618 CGAGAGTACGCGGCTACCTCAGGAGGATCGGCTTCGAGGATCGGCTTCGACTAGCT 677

QY 603 CAAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCGGT 662  
DB 678 CAAGGGCTACGGAGCGTGGGTGCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCGGT 737  
QY 663 TGGAGAGTACTGGGACACCAACCGTCCGACGCTGTTCTCAACTGGGCATATCTCGAGCGGTGC 722  
DB 738 CGGGAGTACTGGGACACCAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 797  
QY 723 CAAGGTCTTTGACTTCGCCCTCTACTACAAGATGGATGAGGCTTTGACAAACAAACAT 782  
DB 798 AAAAGTCTTCGACTTCGCCCTCTACTACAAGATGGAGCGGCGCTTTGACAAACAAACAT 857  
QY 783 TCCAGGCTCGTCTCTGCCCTTCAGAACGCCAGACGTTGTTGCTCCGCGGACCGCTTCAA 842  
DB 858 TCCGCGACTCGTGGGCGCTTCAAGAACGGGGGACAGTCTGAGCGCGGACCGCTTTAA 917  
QY 843 GGCCTTAACCTTTGTAGCAACCAACGACCGATATAATCTGGAAACAAGTATCCAGCTA 902  
DB 918 GGCCTTAACCTTTGTAGCAACCAACGACCGATATAATCTGGAAACAAGTATCCAGCTA 977  
QY 903 CGGTTTCATCTCCTACCTACGAGGGCGGACCAATATCTCCGCGACTACGAGGAGTG 962  
DB 978 CGCCTTCATCTCCTACCTACGAGGGCGGACGATATCTCCGCGACTACGAGGAGTG 1037  
QY 963 GCTCAACAGGATAAGCTCAAGAACCTCATCTGATACATGAGAACCTCCCGGAGGAG 1022  
DB 1038 GCTCAACAGGAGACGAGCTCAAGAACCTCATCTGATACAGCACCTCGCCGCTGGAAG 1097  
QY 1023 CACGACATAGTCTACGATTAACGATGAACCTATCTTCGTCAGGAAACGCTACGGGA 1082  
DB 1098 CACGACATAGTCTACGATTAACGATGAACCTATCTTCGTCAGGAAACGCTACGGGA 1157  
QY 1083 CAAGCCGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGAGGTTGTTA 1142  
DB 1158 CAAGCCGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGAGGTTGTTA 1217  
QY 1143 TGTGCCGAATTCGCGGGCGGTGATCCACAGTATATCTGTAACTCGGAGGCTGGGT 1202  
DB 1218 CGTTCGAAGTTCGCGGGAGCGTGCATCCACAGTATACCGGCAACCTCGGCGCTGGGT 1277  
QY 1203 AGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGC 1262  
DB 1278 GGAAGTGGTGGATCAAGCGGCTGGGTGATCTACCTCGAGGCGCTTCCACGACCCGCG 1337  
QY 1263 CAAGCGGAGTATGGTCTACTCCGTTGGAGCTACTCGGGTGGGCTGA 1311  
DB 1338 CAAGCGGATTTACGGCTACTCCGTTGGAGCTACTCGGGTGGGCTGA 1386

## RESULT 4

AF504063  
LOCUS  
DEFINITION Thermococcus sp. 'AEP11 la' alpha-amylase precursor (BD5031) gene, complete cds.  
ACCESSION AF504063  
VERSION AF504063.1 GI:21326996  
KEYWORDS  
SOURCE ORGANISM Thermococcus sp. 'AEP11 la'  
Thermococcus sp. 'AEP11 la'  
Archaea; Euryarchaeota; Thermococci; Thermococcales;  
Thermococcaceae; Thermococcus.  
1 (bases 1 to 1386)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabelli,M., Lam,D.,  
Macomber,J., Short,J.M., Robertson,D.E. and Miller,C.  
A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE  
J. Biol. Chem. 277 (29), 26501-26507 (2002)  
11994309  
2 (bases 1 to 1386)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabelli,M., Lam,D.,  
Macomber,J., Short,J.M., Robertson,D.E. and Miller,C.  
Direct Submission  
Submitted (18-APR-2002) Bioinformatics, Diversa Corporation, 4955



Directors Place, San Diego, CA 92121, USA  
 Location/Qualifiers  
 1..1386  
 /organism="Thermococcus sp. 'ABPII 1a'"  
 /mol\_type="genomic DNA"  
 /strain="ABPII 1a"  
 /db\_xref="taxon:195101"  
 1..1386  
 /gene="BD5031"  
 1..1386  
 /gene="BD5031"  
 /codon\_start=1  
 /product="alpha-amylase precursor"  
 /protein\_id="AAM48113.1"  
 /db\_xref="GI:21326997"  
 /translation="MKKEVALFITMFVSVMAVVAQASAAKYSELEEGGVIMQAFYV  
 DYPGGIWDITRISKIPWEGYSAIWIPIPAKSMGSGYDYPDYDFDLGYNOK  
 CTIEFRGSKBLINIMTAHAYGIKVADIIVNHRAGDLEWNPFGVDYTFWTFDYSK  
 ASGKYANLDFPHNEKCCDEGTFGGPDIAHEKSDQHLWASDRSYAAYLBSIGV  
 DAMREDYKGVGMVVDLWNGWMAVEYWDNVDALLNWAYSSGAKUPDFPLYK  
 MDEAFDNKIIPALVSALQNGQTVVSRDFKAFVFNHDTDIINWKYLAFILITYEG  
 QVPIPYRDEBWLKNDRLNLIWIHDLAGGSTIVYDSDMIIFVRNGYSGKPLIT  
 YINLGSSKVGWVYVPKPAFCIHYETGNLGVWDKYVYSSGWVYLEAPAYDPANGQY  
 GYSVMSYCGVG"  
 1..78  
 /sig\_peptide  
 1..78  
 /gene="BD5031"  
 ORIGIN  
 Query Match 79.4%; Score 1040.8; DB 1; Length 1386;  
 Best Local Similarity 87.2%; Pred. No. 1.2e-224;  
 Matches 1141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;  
 4 GCCAAGTACTCGAGCTGGAAGAGCGGGCTCATATGACAGGCGTCTTACTGGAGCTG 63  
 Db 79 GCAAGTATTCGAGCTCGAAGAGCGGGCTTATATGACAGGCGTCTTACTGGAGCTG 138  
 64 CTTTCAGGAGGAATATGTTGGACACAATACGACAGAGATACCGAGTGTGACGATGCC 123  
 Db 139 CCAGGTGGAGGAATCTGTTGGGACACCATCAGGAGCAAGATACCGAGTGTGACGAGCG 198  
 124 GGAATCTCGCAATATGATTTCCCGCGGAGACAGGACATGGCGGGCGCTTATCGATG 183  
 Db 199 GGAATATCGGCATTTGATTTCCGACGACGACGAGGGGATGAGCGGGGTACTCGATG 258  
 184 GGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTAGAG 243  
 Db 259 GGTACGATCCCTACGATTTCTTTGACCTCGCGAGTACACACAGAGGAAACCATCGAA 318  
 244 AGCGCTTTGGCTCCAGCAGAGCTCTGTGAACATGATATAACACCGCCACGCTATGGC 303  
 Db 319 AGCGCTTTGGCTCTAAACAGAGCTCATCAATATGATATAACACCGCCACGCTATGGC 378  
 304 ATGAAGGTAATAGCGGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGGAC 363  
 Db 379 ATAAAGGTCATAGCGGACATGTCATAAACACCGCGCGCGGAGACCTCGAGTGGAC 438  
 364 CCCTTCGTGAACGACTATACCTGGACCGGACTTCTCAAAAGGTGCGGTGCGGTGAATACAG 423  
 Db 439 CGCTTCGTTGGGACTACACCTGGACCGGACTTCTCAAAAGGTGCGGTGCGGTGAATACT 498  
 424 GCCAACTACCTCGACTTCCACCCGACAGCTCATCGATCGCGGCGATTCGGAACATTTGGA 483  
 Db 499 GCCAACTACCTCGACTTCCACCCGACAGCTCATCGATCGCGGCGGACATTTGGA 558  
 484 GGCTATCCGACATATGCCACAGAGCTGGGACCACTGCTCTCGGCGGACGACG 543  
 Db 559 GGCTTCCAGACATAGCCACGAGAGAGCTGGGACCACTGCTCTGGGCGAGGAT 618  
 544 GAGAGCTACGCGGACATATCTCAGGAGCATTCGCGATTCGATCGCTCGGCTTCGATACGTC 603  
 Db 619 GAGAGCTACGCGGCTTACCTAAGGAGCATCGCGCTTATGCTCGGCGCTTGTGACTAGCTG 678

604 AAGGGCTATGCTCCCTGGGTGCTCAAGGAGTGGCTGAACTGGTGGGAGAGCTGGGCGGTT 663  
 Db 679 AAGGGCTACGAGAGCTGGGTGCTCAAGGAGTGGCTGAACTGGTGGGCGGCTGGGCGGTT 738  
 664 GGAGAGTACTGGGACACCAACGTCGAGCGTGTCTCAACTGGGCATACTCGAGCGGTGCC 723  
 Db 739 GCGGAGTACTGGGACACCAACGTTGATGACTCCTCAACTGGGCTTACTCGAGCGGCGCC 798  
 724 AAGGTCTTGTGACTTCGCCCTCTACTCAAGATGGATGAGGCGCTTTCACAAAAACATT 783  
 Db 799 AAGGTCTTGTGACTTCGCCCTCTACTCAAGATGGATGAGGCGCTTTCACAAAAACATT 858  
 784 CCAGCGCTGCTCTGCCCTTCAGAGCGGCGGAGCTGTTGCTCCGCGGACCGTTCAG 843  
 Db 859 CCAGCGCTGCTCTGCCCTTCAGAGCGGCGGAGCTGTTGCTCCGCGGACCGTTCAG 918  
 844 GCGGTAACTTTCTAGCAAAACCAACGACACCGATATATCTGAAACAAAGTATCCAGGCTAC 903  
 Db 919 GCGGTAACTTTCTAGCAAAACCAACGACACCGATATATCTGAAACAAAGTATCCATTGTTAT 978  
 904 GCGTTCATCTCACTACGAGGCGGCGGACCAATATTTCTACCGGACTACGAGGAGTGG 963  
 Db 979 GCTTTTCATCTCACTACGAGGCGGCGGACCAATATTTCTACCGGACTACGAGGAGTGG 1038  
 964 CTCAACAGGATTAAGCTCAAGAACCTCATCTGATATACAGAACCTCGCGGAGGAAGC 1023  
 Db 1039 CTCAACAGGATTAAGCTCAAGAACCTCATCTGATATACAGAACCTCGCGGAGGAAGC 1098  
 1024 ACCGACATAGTCTACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1083  
 Db 1099 ACCGACATAGTCTACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1158  
 1084 AAGCGGGGCTTATACCTACATCAACCTAGCTGAGCAAGCGGCGGAGGTTGGTTTAT 1143  
 Db 1159 AAGCTTGGCTTATACCTACATCAACCTAGCTGAGCAAGCGGCGGAGGTTGGTTTAT 1218  
 1144 GTCCGGAAGTTCCGCGGCGGCTGATCCAGAGTACTTGGTAACTCCGAGGCTCGGTA 1203  
 Db 1219 GTCCGGAAGTTCCGCGGCGGCTGATCCAGAGTACTTGGTAACTCCGAGGCTCGGTA 1278  
 1204 GACAGTACGCTTACTCAAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
 Db 1279 GACAGTACGCTTACTCAAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1338  
 1264 AAGCGGAGTATGGCTACTCCGCTGAGCTACTCGGCGGTGGGCTGA 1311  
 Db 1339 AAGCGGAGTATGGCTACTCCGCTGAGCTACTCGGCGGTGGGCTGA 1386

RESULT 5  
 AF068255  
 LOCUS  
 DEFINITION  
 Thermococcus hydrothermalis strain AL662 alpha-amylase (amy) gene,  
 complete cds.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Thermococcus hydrothermalis  
 Thermococcus hydrothermalis  
 Archaea: Euryarchaeota; Thermococci; Thermococcales;  
 Thermococcaceae; Thermococcus.  
 REFERENCE  
 1 (bases 1 to 2705)  
 AUTHORS  
 Leveque, E., Belarbi, A. and Haye, B.  
 JOURNAL  
 Patent: France 98.05655 05-MAY-1998;  
 REFERENCE  
 2 (bases 1 to 2705)  
 AUTHORS  
 Leveque, E., Haye, B. and Belarbi, A.  
 TITLE  
 Cloning and expression of an alpha-amylase encoding gene from the  
 hyperthermophilic archaeobacterium Thermococcus hydrothermalis and  
 biochemical characterisation of the recombinant enzyme  
 JOURNAL  
 FEMS Microbiol. Lett. 186 (1), 67-71 (2000)  
 MEDLINE  
 20243462  
 PUBMED  
 10779714  
 REFERENCE  
 3 (bases 1 to 2705)

AUTHORS Leveque, E., Hays, B. and Belarbi, A.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAY-1998) Laboratory of General and Molecular Microbiology, UPR Sciences, University of Reims-Champagne-Ardenne, Moulin de la Housse, BP 1039, Reims Cedex 2 51687, France

## FEATURES

Location/Qualifiers  
 1..2705  
 /organism="thermococcus hydrothermalis"  
 /mol\_type="genomic DNA"  
 /strain="AL662"  
 /db\_xref="taxon:46539"  
 424..1797  
 /gene="amy"  
 424..1797  
 /gene="amy"  
 /EC\_number="3.2.1.1."  
 /codon\_start=1  
 /transl\_table=11  
 /product="alpha-amylase"  
 /protein\_id="AAC97877.1"  
 /db\_xref="GI:4049918"  
 /translation="MARKVIVALLPLVLVSVAIPAKAETLENGGVIMQAFYWDVPG  
 GSIWDTIAQKIPDASAGISAIWIPSPKSMGSGYDYPDFDILGEYKQSGVE  
 TFSGKEBLVNTIAHANMKVIADIVINHRAGDLEWNPFTSYTWTDFSKVASGK  
 YTANYLDFHPLNLAGDSFGYDPICHDKSWDQHLWASNESYAAVLRISIGIDAWR  
 EDVKGYAPVVKWLNRRGWAGVEYDNTVDALLSWAYDSKAFVDFPLVYKMDER  
 FNNNTPALVDALKNKGTVVSDPDKATVFVANHDNTIINWKPAYAPVILYVGGPAI  
 FRDYEWLNKRLNLIWHDLHLAGGSDIILYDSDELIIVRNGYVGRKPLIIFYINL  
 GSKAGRWVYVPRFAGSCIHEYITGNLGGWIDRWVSSGRVYLEAPHPANGQYISV  
 WSYCGVG"

## ORIGIN

Query Match 76.5%; Score 1003.4; DB 1; Length 2705;  
 Best Local Similarity 86.0%; Pred. No. 3 4e-216;  
 Matches 1112; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

19 CTGAAAGGCGGGTCTAATGAGGCGTCTCTAGGAGCGTGCCTTCAGAGGAATA 78  
 505 CTTGAGAACGGCGGCTCATAATGAGGCGTCTCTAGGAGCGTGCCTTCAGAGGAATA 564  
 79 TGGTGGGACACATACGGGAGAGATACCGGAGTGTACGATGCGGGAATCTCCGCAATA 138  
 565 TGGTGGGACACATACCGGAGAGATACCGGAGTGTACGATGCGGGAATCTCCGCAATA 624  
 139 TGGATTCCTCCCGGAGTACGAGGCGGCGGCGCTATTTCGATGGGCTACGACCCCTAC 198  
 625 TGGATTCCTCCCGGAGTACGAGGCGGCGGCGCTATTTCGATGGGCTACGACCCCTAC 684  
 199 GACTTCCTTCGAGTACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 258  
 685 GATTTCCTTCGAGTACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 744  
 259 AAGCAGGAGCTCTGGAACATGATAACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 318  
 745 AAGCAGGAGCTCTGGAACATGATAACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 804  
 319 GATATAGTATCAACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 378  
 805 GATATAGTATCAACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 864  
 379 TATACCTGGACCGGCTCTCAAGGTCGCGTCCGCGGTAATACACGCGGCGGCGGCGGCGG 438  
 865 TACCTGGACCGGCTCTCAAGGTCGCGTCCGCGGTAATACACGCGGCGGCGGCGGCGGCGG 924  
 439 TTCCACCGGACGAGCTCTCAAGGTCGCGTCCGCGGTAATACACGCGGCGGCGGCGGCGG 498  
 925 TTCCACCGGACGAGCTCTCAAGGTCGCGTCCGCGGTAATACACGCGGCGGCGGCGGCGG 984  
 499 TGCACACGAGAGCTCTCAAGGTCGCGTCCGCGGTAATACACGCGGCGGCGGCGGCGGCGG 558  
 985 TGCACACGAGAGCTCTCAAGGTCGCGTCCGCGGTAATACACGCGGCGGCGGCGGCGGCGG 1044  
 559 TATCTCAGGAGCATCGGCATCGATCGCTCGGCGGCTTCGACTACGTCAGGCGCTATGCTCCC 618

1045 TACCTCCGAGGATCGGCATCGACGCTGGCGCTTCGACTACGTCAGGCTACGCTCCC 1104  
 619 TGGGTCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCGGTTGGAGAGTACTGGGAC 678  
 1105 TGGGTCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCGGTTGGAGAGTACTGGGAC 1164  
 679 ACCAAGCTCGACGCTGTTCTCACTCGGCATATCTCGAGCGGTGCCAAGTCTTTGACTTC 738  
 1165 ACCAAGCTCGATGCTCTCTGAGCTGGGCTACGACGCGGTCTAAAGTCTTCGACTTC 1224  
 739 GCCCTCTACTCAAGATGGATGAGGCTTTTGACAAACAAACAAANTCCAGCGCTGCTCT 798  
 1225 CCGCTCTACTCAAGATGGAGGAGGCTTCGATAACAAACATCCCGCGCTCGTGGAC 1284  
 799 GCCCTTCAGAACCGGACGACTGTTCTCCCGGAGCGGCTTCAGGCGGCTTAACCTTTGA 858  
 1285 GCCCTCAAGAACCGGAGGCGGCTGCTGAGCGGCGGCGGCTTCAGGCGGCTGCTCTG 1344  
 859 GCAAAACACGACCGGATATAATCTCGAAACAAGTATCCAGCCTACGCTTCATCTCACC 918  
 1345 GCAAAACACGATACCAACAATACTCGAAACAAGTATCCAGCCTACGCTTCATCTCACC 1404  
 919 TACGAGGCGGAGCGGACCAATATCTACCGGAGCTACGAGGAGTGGCTCAACAGGATAAG 978  
 1405 TATGAGGAGACGCGGCAATATCTACCGGAGCTACGAGGAGTGGCTCAACAGGAGCAGG 1464  
 979 CTCAAGAACCTCTCTCGATACATGAGAACCTCCCGGAGGAGGAGCAGCAGATAGTCTAC 1038  
 1465 CTCAAGAACCTCTCTCGATACATGAGAACCTCCCGGAGGAGGAGCAGCAGATAGTCTAC 1524  
 1039 TACGATAACGATGAATCTCTCTCGTCAAGAACGCTTACCGGAGCAGGCGGCTTATA 1098  
 1525 TACGACGAGCAGGAGCTTATCTCTCGTGAAGAACGCTTACCGGAGCAGGCGGAGTATA 1584  
 1099 ACCTACATCAACCTTAGGCTCGAGCAGGCGGAGGAGTGGGCTTTATGTCGCGAAGTTCGCG 1158  
 1585 ACCTACATCAACCTTAGGCTCGAGCAGGCGGAGGAGTGGGCTTTATGTCGCGAAGTTCGCG 1644  
 1159 GCGCGCTGCATCCACGAGTATCTGTAACCTCGGAGGCTGGGTAGACAGAGTACGCTAC 1218  
 1645 GGCTCGTGATACACGAGTACACCGGCAACCTCGGCGGCTGGATTCACAAAGTGGGTGAC 1704  
 1219 TCAAGCGGCTGGGTCTATCTCGAGCTTCAGGCTTACGACCTCCGAGCGGAGTATGGC 1278  
 1705 TCAAGCGGCTGGGTCTATCTCGAGCTTACGACCTCCGAGCGGAGTATGGC 1764  
 1279 TACTCCGTGGAGCTACTCGGCGGCTGGGCTGA 1311  
 1765 TACTCCGTGGAGCTACTCGGCGGCTGGGCTGA 1797

## RESULT 6

D83793

LOCUS

D83793

DEFINITION

Pyrococcus sp. DNA for alpha-amylase, complete cds.

ACCESSION

D83793

VERSION

D83793.1

KEYWORDS

Apka; alpha-amylase.

SOURCE

Pyrococcus sp.

ORGANISM

Pyrococcus sp.

Archaea; Euryarchaeota; Thermococci; Thermococcales;

1 (sites)

Tachibana, Y., Mendez, L., Fujiwara, S., Takagi, M. and Imanaka, T.

cloning and expression of the alpha-amylase gene from the

hyperthermophilic archaeon Pyrococcus sp. KOD1, and

characterization of the enzyme

J. Ferment. Bioeng. 82, 224-232 (1996)

2 (bases 1 to 2179)

Tachibana, Y., Mendez, L., Takagi, M. and Imanaka, T.

Direct Submission

Submitted (05-MAR-1996) Yoshihisa Tachibana, Osaka University,

Faculty of Engineering, Biotechnology; 2-1, Yamadaoka, Suita, Osaka  
565, Japan (Tel:06-879-7442, Fax:06-879-7448)

## FEATURES

Location/Qualifiers  
1..2179  
/organism="Pyrococcus sp."  
/mol\_type="genomic DNA"  
/strain="KOD1"  
/db\_xref="taxon:33866"  
452..1837  
/gene="Apka"  
452..1837  
/gene="Apka"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-amylase"  
/protein\_id="BAA21130.1"  
/db\_xref="GI:2251108"  
/translation="MKKFVLLITMPVVSMAVAOPASAAKYSELEGGVIMQAPYV  
DVPAGGIWDTIRSKIPWEYEGISALWIPPSKMGXGAYSGDYDPDPLGEYNQK  
GVETREGSKQELINHTAHYGIKVIADIVINHRAGGLEWNPVFGYITWTFPSKV  
ASQKYTYLDFHPNEVKCDEGTFGGFDPIAHKEKWDQHLWASDEKAYLRSIGV  
DAWFDYVKGAWVVDLWNGWAGVYDNTVDLNNWAYSGAKVDFPPLYK  
MDAFDNTNIPALVDALQNGTVVSRDPFKAVTFVANHDTDIWNKYPAYAFILITVEG  
QPIFYRDEBEWLKDKLNLINHDLHAGSSTIVYDSDELIFVRNGDSKRPGLIT  
YINLSSKVRGVVYVPKFAAGTCEHYTGNLGGVVDKVVESSGWVYLEAPAYDPASQY  
GTVNSYCGVG"  
452..529  
/gene="Apka"  
530..1834  
/gene="Apka"  
/product="alpha-amylase"

## sig\_peptide

## mat\_peptide

## ORIGIN

Query Match 75.6%; Score 991.2; DB 1; Length 2179;  
Best Local Similarity 84.9%; Pred. No. 2e-213;  
Matches 1110; Conservative 0; Mismatches 198; Indels 0; Gaps 0;  
4 GCCAAGTACTCCGAGTCGGAAGGCGGGTCTAATGTCAGCGGTCTCTACTGGAGCTG 63  
530 GCAAGTATCCGAACCTCGAAGAGGGCGGCTTATAATGCAAGGCTTCTACTGGAGCTT 589  
64 CTTTCAGGAGGAATATGTTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGATGCC 123  
590 CCACGGGAGGAATCTGGTGGAATACATCAGAGCAAGATACCGGAGTGGTACGAGGCT 649  
124 GGAATCTCCGAATATGATATCCCGCGGAGGAGGCAATGGCGGCGCTTATTCGATG 193  
650 GGAATCTCCGCAATCTGATTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 709  
184 GGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTAGAG 243  
710 GGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTAGAG 769  
244 ACGGCTTTGGCTCAGAGCAGGAGCTCGTGAACTATGAAATACACCGCCACGCTATGCG 303  
770 ACTCGCTTTGGCTCAAAGCAGGAGCTTATCAACATGATATAACACCGCCACGCTATGCG 829  
304 ATGAAGGTATAGCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTGAAC 363  
830 ATAAAGGTATAGCTGATATCGTATATAACCGCGCGCGGAGACCTCGAGTGAAC 889  
364 CCCTTCGTGAACGACTATACCTGACCGGACTTCTCAAAGGTGCGGTGCGGTAAATACAG 423  
890 CCGTTGTTGGGACTACACCTGACCGGACTTCTCAAAGGTGCGGTGCGGTAAATACAG 949  
424 GCCAATCTCGACTTCACCGGAAACGAGCTCCATCGGGCGGATTCGGAGACATTTGGA 493  
950 GCCAATCTCGACTTCACCGGAAACGAGCTCCATCGGGCGGATTCGGAGACATTTGGA 1009  
484 GGCTATCCCGACATATGCGACGAGAGCTGGAGACGAGTACTGGCTCTGGGCGGAGCAG 543  
1010 GGTTCACGACATAGCCACGAGAGGAGTGGAGCAGCACTGGCTCTGGGCGGAGCAG 1069

544 GAGAGCTACGGGCGCATATCTCAGGAGCATCGGCATCGATGCTGCGGCTTCGACTACGTC 603  
1070 GAGAGCTACGGGCGCATCTCAGGAGCATCGGCATCGATGCTGCGGCTTCGACTACGTC 1129  
604 AAGGCTATGCTCCTGGTGGTCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGCGGT 663  
1130 AAGGCTACGGAGCTGGTGGTCAAGGACTGGCTCAAACTGGTGGGCGGCTGGGCGGTC 1189  
664 GGAGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCACTACTCGAGCGGTGCC 723  
1190 GGTGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCACTACTCGAGCGGCGCC 1249  
724 AAGTCTTTGACTTCGCGCTCTACTCAAGATGATGAGGCTTTGACAAACAAACATTT 783  
1250 AAGTCTTTGACTTCGCGCTCTACTCAAAATGGACGAGGCTTTGACAAACAAACATTT 1309  
784 CCAGCGCTGCTCTGCGCTTCAGAAACGCGCAGACTGTGTCTCCGCGACCCGCTTCAAG 843  
1310 CCGGCTTGGTGTGATGCGCTTCAGAAACGCGGGAACCGTCTCTCGGCGACCGTTCAAG 1369  
844 GCGTAAACCTTTGTAGCAAAACCAACACCGCATATAATCTGGAAACAATATTCACGCTAC 903  
1370 GCGTAAACCTTTGTAGCAAAACCAACACCGCATATAATCTGGAAACAATATTCACGCTAC 1429  
904 GCGTTCATCCTCACCTACGAGGCGCGCAGCAATATCTACCGGACTACGAGGAGTGG 963  
1430 GCTTTTATCTCCTACCTACGAGGCGCGCGCTCATATTTCTACCGGACTACGAGGAGTGG 1489  
964 CTCAACAGGATTAAGCTCAAGAAACCTCATCTGGATACATGAGAACTTCGTCAGGAGGAAAGC 1023  
1490 CTCAACAGGATTAAGCTTAACAACTAATCTGATACACGACCACTCGCGGTTGGAAGC 1549  
1024 ACGCATAGTCTACTACGATTAACGATGAATCTATCTGTCAGGAAACGGCTACGGGAC 1083  
1550 ACGCATAGTCTACTACGATTAACGATGAATCTATCTGTCAGGAAACGGCTACGGGAC 1609  
1084 AAGCGGGGCTTATTAACCTACATCACTAGCTAGCTCGAGCAAGCGCGGAGTGGGTTTAT 1143  
1610 AAGCGGGGCTTATTAACCTACATCACTAGCTAGCTCGAGCAAGCGCGGAGTGGGTTTAT 1669  
1144 GTGCGGAATTCGCGGCGGCTGTCATCCAGATGATCTGTTAACTCGAGGCTTCGAGGCTGGGTA 1203  
1670 GTGCGGAATTCGCGGCGGCTGTCATCCAGATGATCTGTTAACTCGAGGCTTCGAGGCTGGGTA 1729  
1204 GACAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
1730 GACAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1789  
1264 AACGGGCACTATGGCTACTTCGCTGTGGAGCTACTCGGGGTGGGCTGA 1311  
1790 ACGGCGAGTACGGCTACACCGTCTGGAGCTACTCGGGGTGGGATGA 1837

## RESULT 7

E13334  
LOCUS  
DEFINITION  
GDNA encoding alpha-amylase.  
E13334  
ACCESSION  
E13334  
VERSION  
E13334.1 GI:3252139  
KEYWORDS  
JP1997173077-A/1.  
SOURCE  
unidentified  
ORGANISM  
unclassified.  
REFERENCE  
1 (bases 1 to 2179)  
AUTHORS  
Imanaka, T., Tachibana, Y., Suzuki, Y., Kojima, I. and Uzura, K.  
TITLE  
ULTRATHERMOSTABLE ACIDIC ALPHA-AMYLASE AND DNA FRAGMENT  
CONTAINING GENE FOR PRODUCING THE ALPHA-AMYLASE  
JOURNAL  
Patent: JP 1997173077-A 1 08-JUL-1997;  
IMANAKA TADAYUKI, NAGASE SEIKAGAKU KOGYO KK  
COMMENT  
OS Unclassified  
PN JP 1997173077-A/1  
PD 08-JUL-1997  
PF 19-JUL-1996 JP 1996191138

```

PR 20-JUL-1995 JP 95P 216454
PI IMANAKA TADAYUKI, TACHIBANA YOSHINAGA, SUZUKI YUJI, PI
KOTJIMA IWAO,
PC C12N15/09, C07H21/04, C12N1/21, C12N9/30, C12P19/14, (C12N15/09, PC
C12R1:01),
CC (C12N1/21, C12R1:19), (C12N9/30, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2179
FT /organism='Unclassified'
FT /strain='KOD-1'
FT CDS 1<.2179
FT /product='alpha-amylase'.
FT Location/Qualifiers
1..2179
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
ORIGIN
Query Match 75.5%; Score 989.6; DB 6; Length 2179;
Best Local Similarity 84.8%; Pred. No. 4.5e-213;
Matches 1109; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 4 GCGAGTACTCGAGCTGGAAGGGGGGTCATTAATGCAGGCGTTCTACTGGAGGTG 63
DB 530 GCAAAAGTATTCGCACTCGAAGAGCGCGGCTTAAATGCAGGCGTTCTACTGGAGCGTT 589
QY 64 CTTTCAGGAGGATATGTGGGACACATACCGGAGAGATACCGAGTGGTACGATGCC 123
DB 590 CACGGGAGGATCTGGTGGATACATCAGAACAGATACCGAGTGGTACGAGGCT 649
QY 124 GGAATCTCGGATATGATATCCCGGCGAGCAAGGGGATGGCGGCGCTTATTCGATG 183
DB 650 GGAATCTCGGCGATCTGGATTCGCGAGCCAGCAAGGGGATGGGAGGAGCTTATTCATG 709
QY 184 GCTAGGACCCCTAGACCTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGTTAGAG 243
DB 710 GGTAGACCCATAGACCTCTTCGACCTCGGCGAGTACACCAAGAGGAAACGTTGAA 769
QY 244 ACGCGCTTTGGCTCCAGCAGAGCTCGTGAAATGATAAACAACCGCCACGCTATGGC 303
DB 770 ACTCGCTTTGGCTCAAAGCAGGAGCTTATCAACATGATAAACAACCGCCCTACGCG 829
QY 304 ATGAGGTATACCGATATAGTATCAACACCGCGCGGCTGACCTGGAGTGGAGC 363
DB 830 ATAAAGGTATAGTATAGTATCGTATAAACCAACCGCGCGGCGGAGCTCGAGTGGAA 889
QY 364 CCGTCTGTGAACGACTATACCTGGAGCGGACTTCTCAAAGGTGCGCTCGGGTAAATACAG 423
DB 890 CCGTCTGTGGGACTACACCTGGAGCGGACTTCTCGAAGGTGGCTCGGGCAATATACC 949
QY 424 GCGAATAGTCTGACTTCCACCGCAAGAGCTCCATGCGGGGATTCGGAACATTGGA 483
DB 950 GCGAATAGTCTGACTTCCACCGCAAGAGCTCCATGCGGGGATTCGGAACATTGGA 1009
QY 484 GCGTATCCGACATATGCCACACAGAGCTGGGACCACTGCTGGCTCGGGCCAGCCAG 543
DB 1010 GGTTCACAGATAGCCACAGAGGAGGAGGACGACTGGCTCTGGCGAGCGAC 1069
QY 544 GAGAGCTACCGGCAATATCTCAGAGGATTCGAGCATGCTGCGGCTTCGATACGTC 603
DB 1070 GAGAGCTACCGGCGCTACCTCAGAGGATTCGCGGCTTGTGCTGCGGCTTCGATACGTA 1129
QY 604 AAGGGCTATGCTCCCTGGGCTCGTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGGTT 663
DB 1130 AAGGGCTACGAGGCTGGGCTCGTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGGCT 1189

```

```

QY 664 GGAGAGTACTGGGACACCAACGCTCGAGCGTGTCTCACTGGGCGATCTCGAGCGGTGCC 723
DB 1190 GGTGAGTACTGGGACACCAACGCTGTGATGACCTCTCAATTTGGGCGATCTCGAGCGGTGCC 1249
QY 724 AAGGTCTTTGACTTCCGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAACATT 783
DB 1250 AAGGTCTTGGACTTCCGCTCTACTCAAAATGGAGGAGGCTTTGACAAACAAACATT 1309
QY 784 CAGAGCGCTGTCTCTCGGCTTCAGAGCGCCAGAGCTGTGTCTCCGCGGACCGCTTCAAG 843
DB 1310 CCGGCTTGGTGTGATGCGCTTCAGAAACCGGGGAAACCGTGTCTCTCGGACCGCTTCAAG 1369
QY 844 GCGTAAACCTTTGTAGCAAAACCAACGACCGGATATAATCTGGAACAAGTATCCAGCTAC 903
DB 1370 GCGTAAACCTTTGTAGCAAAACCAACGACCGGATATAATCTGGAACAAGTATCCAGCTAC 1429
QY 904 GCGTAAACCTTTGTAGCAAAACCAACGACCGGATATAATCTGGAACAAGTATCCAGCTAC 963
DB 1430 GCGTAAACCTTTGTAGCAAAACCAACGACCGGATATAATCTGGAACAAGTATCCAGCTAC 1489
QY 964 CTCAACAGGATTAAGCTCAAGAAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGAC 1023
DB 1490 CTCAACAGGATTAAGCTCAAGAAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGAC 1549
QY 1024 ACCGATAGTCTTACGATACGATGAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGAC 1083
DB 1550 ACCGATAGTCTTACGATACGATGAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGAC 1609
QY 1084 AAGCGGGGCTTATTAACCTTACATCAACCTAGGCTCGAGCAAGCGCGGAGGAGGTTTAT 1143
DB 1610 AAGCGGGGCTTATTAACCTTACATCAACCTAGGCTCGAGCAAGCGCGGAGGAGGTTTAT 1669
QY 1144 GTCCCGAAGTTCGCGGCGCTGCACTCCACAGTATCTGTAACCTCGAGAGGCTGGGTA 1203
DB 1670 GTCCCGAAGTTCGCGGCGCTGCACTCCACAGTATCTGTAACCTCGAGAGGCTGGGTA 1729
QY 1204 GACAGTACTGCTTACTCAAGCGGCTGGGCTCTATCTGAGGCTCCAGCTTACGACCGCTGCC 1263
DB 1730 GACAGTACTGCTTACTCAAGCGGCTGGGCTCTATCTGAGGCTCCAGCTTACGACCGCTGCC 1789
QY 1264 AAGCGGAGTATGGCTTACTCGCTGTGGAGCTACTGCGGGGTGGGCTGA 1311
DB 1790 AAGCGGAGTATGGCTTACTCGCTGTGGAGCTACTGCGGGGTGGGCTGA 1837

```

```

RESULT 8
AF017454 2372 bp DNA linear BCT 25-JUN-2001
LOCUS Thermococcus sp. Rt3 amylase (amy) gene, complete cds.
DEFINITION Thermococcus sp. Rt3 amylase (amy) gene, complete cds.
ACCESSION AF017454
VERSION AF017454.1 GI:2655009
SOURCE Thermococcus sp. Rt3
ORGANISM Thermococcus sp. Rt3
REFERENCE 1 (bases 1 to 2372)
AUTHORS Jones, R.A., Jermin, L.S., Eastal, S., Patel, B.K. and Beacham, I.R.
TITLE Amylase and 16S rRNA genes from a hyperthermophilic archaeobacterium
JOURNAL J. Appl. Microbiol. 86 (1), 93-107 (1999)
MEDLINE 99154282
PUBMED 10030014
REFERENCE 2 (bases 1 to 2372)
AUTHORS Jones, R.A., Patel, B. and Beacham, I.R.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1997) School of Biomedec. and Biomedical Science,
Griffith University, Nathan, Brisbane, Qld. 4111, Australia
FEATURES
1..2372
/organism="Thermococcus sp. Rt3"
/mol_type="genomic DNA"
/strain="Rt3"
/db_xref="taxon:65421"

```

```

gene      372..1781
/ gene="amy"
CDS       372..1781
/ gene="amy"
/EC_number="3.2.1.1"
/codon_start=1
/transl_table=11
/product="amylase"
/protein_id="AAB87860.1"
/db_xref="GI:2655010"
/translation="MSAKKLLALLFVLAVGVAVIPARVGIAPVSAGATSRPSLEEG
GVIMQAFVDPAGIWDITRSKIPDWASAGISALWIPFASKGMSGISMGIDPYDF
FDLGEYKQGVETRFSGKQLINIMINTAHISYGIKVIADIVINHRAGDLEWNPFTNS
YTWDFSKVAGTFRGKYNLDHPNEVKGCCDEGTGFPDIAHEKSNQYWLWASQKSY
AAYLRSIGIDAWRFDYKGYGAWYKDLKWLALVGEYWDITNVDALLINWAYSSGAKV
FDPLYKMDAFONKNIIPALVSLQNGQTVSRDPFKAVTFVANHDTDIIWNKYPAY
AFILTYGQPIVIFRVDVEEMLNKLRLNLIWHNNLAGGSTIVVYVNDLILFVRNGY
GNKPLITVILNGSKVGRVYVVPKFAAGSIHETGNLGSWDXVYVSGNWVLEAPA
HDPKAGQGYGVSWYICGVG"

ORIGIN
Query Match      73.5%; Score 963.6; DB 1; Length 2372;
Best Local Similarity 84.4%; Pred. No. 3.4e-207;
Matches 1096; Conservative 0; Mismatches 199; Indels 3; Gaps 1;

QY 14 CCGAGCTGAAAGCGCGGCTCATATGCAAGGCTTCTACTGGAGCTGCGCTTCAGGAG 73
Db 487 CCAGCTCGAAGAGCGCGGCTCATATGCAAGGCTTCTACTGGAGCTGCGCTTCAGGAG 546

QY 74 GAATATGTTGGGACACAAATACGGCAGAGATACCGGAGTGTGATCGATGCCGGGAATCTCCG 133
Db 547 GAATCTGTGGATACATATAGGAGCAAGATACCGGACTGGCGCAGCGCTGGAATCTCAG 606

QY 134 CAATATGGAATCCCGCGCAGCAAGGCATGGCGCGCTTATTCGATGGGCTTACGACC 193
Db 607 CTATATGGAATCCCTCCCGCAGCAAGGGTATAGTGGTGGCTTATTCATGGGCTTACGACC 666

QY 194 CCTAGCACTTTTACCTCGTGGTACGACCAAGGGAACGGTAGAGAGCGGCTTTG 253
Db 667 CCTAGCACTTTTACCTCGGCGAGTACTACCAAGAGGGAACCGTTGAGACCGCGCTTTG 726

QY 254 GCTCAAGCAGGAGCTCGTGAACATGATAAACACCGCCACGCTATGCGATGAAGTAA 313
Db 727 GCTCAAGCAGGAGCTAATTAACATGATAAACACCGCCACCTCTCTATGGAATCAAGTCA 786

QY 314 TAGCGGATATAGTANTCAACACCGCGCGCGGTGACTGGAGTGAAACCCCTTCTGTA 373
Db 787 TAGCGGACATAGTCATAAACACCGCGCGCGCGACTTGGAGTGGAAACCCCTTCAAA 846

QY 374 ACGACTATACCTGGACCGACTTCTCAAAGGTGCGGTGCGGTAAATACAGGCCCAACTACC 433
Db 847 ACAGCTACACCTGGACTGACTTCTCAAAGGTGCGTCCGGCAGTACACCGCCAACTACC 906

QY 434 TCGACTTCCACCGCAACGAGTCCATGCGGCGATTCGGGAACATTTGGAGGCTATCCCG 493
Db 907 TCGACTTTCACCCCAACAGAGTCAAAATGCTGTGATGAAGGCACCTTCGAGGGGTTCGCG 966

QY 494 ACATATGCCACGACAGAGCTGGGACGAGTACTGCTCTGGCGCAGCCAGCGAGGAGCTACG 553
Db 967 ACATAGCCCATGAGAAGAGCTGGGACCGAGTACTGCTCTGGCGCAGCCAGAGGAGCTATG 1026

QY 554 CGGCATATCTCAGGAGCATCGGCATCGATGCTGCGCTTCGACTAGCTCAAGGCGCTATG 613
Db 1027 CTGCGCTACCTCAGGAGCATAGGCATCGATGCTGCGCTTCGACTAGCTCAAGGCTTACG 1086

QY 614 CTCCTTGGGTGCTCAGAGACTGGCTGAAGTGGTGGGAGGCTGGCGGTTGGAGAGTACT 673
Db 1087 GGGCGTGGGTGGTCAAGGACTGGCTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1143

QY 674 GGGACACCAAGCTGACGCTGTTCTCAACTGGGATACCTGAGCGGCTGCGCAAGGTCTTTG 733
Db 1144 GGGATACAAACGTTGATGCCCTCTTAAGTGGGCTACTCAAGCGGTGGCCAGGTCTTCG 1203

```

```

QY 734 ACTTCGCCCTCTACTACAAGATGGATGAGCGCTTTTGACAAACAAACATTCAGCGCTCG 793
Db 1204 ACTTCGCCCTCTACTACAAGATGGATGAGCGCTTTTGACAAACAAACATTCAGCGCTCG 1263

QY 794 TCTTCGCCCTCTCAGAACCGCCAGACGTGTTGTCTCCGCGACCCGTTCAAGGCGGTAACT 853
Db 1264 TTTTCGCCCTCTCAGAACCGCTCAGACAGTGTGTTTCCGCGACCCCTTCAAGGCGGTAACT 1323

QY 854 TTTTCAGAAACACGACGACCGCATATATCTGGAAACAAGTATCCAGCTTACGGCTTCATCC 913
Db 1324 TGTTCGCCAAACACGACGACCGCATATATCTGGAAACAAGTATCCGGCTTATGCTTCATCC 1383

QY 914 TCACCTACGAGGCCACGCGCAACAATTTACCGGACTACGAGGAGTGGCTCAACAAG 973
Db 1384 TCACCTACGAGGCTCAGCGGTTATATTTCTACCGGACTACGAGGAGTGGCTCAACAAG 1443

QY 974 ATAAAGCTCAAGACCTCATCTGGATACATGAACTCTCCGCGGAGGAACGACGACATAG 1033
Db 1444 ACAGACTCAAGAACCTCATCTGGATACACAACAATCTCCGCGCGGAGCAGCAGCATAG 1503

QY 1034 TCTTACTACGATACGATGAACCTCATCTTCTCAGGAACCGGCTACGCGGAGCAAGCCCGGCG 1093
Db 1504 TCTTACTACGATACGAGGAGCTCATCTTCTCAGGAACCGGCTACGCGGAGCAAGCCCGGCG 1563

QY 1094 TTATAACCTTACATCAACCTAGGCTCGAGCAAGCCCGAAGGTGGGTTTATGTGCCGAAGT 1153
Db 1564 TTATAACCTTACATCAACCTCGGCTCCAGCAAGCTCGAAGGTGGGTTTATGTGCCGAAGT 1623

QY 1154 TCGCGGCGCGTGCATCCACAGTATACCTCGGAGGCTGGGTAGCAAGTACG 1213
Db 1624 TCGCGGCTTCTGTCATCCACAGTACACCGGCAACCTCGCGGCTGGGTAGCAAGTACG 1683

QY 1214 TCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCAGACTTACGACCTGCAACGCGGCACT 1273
Db 1684 TTTGGCTCAAAACGGATGGGTCTACCTCGAGGCGCTGCCACGACCTGCAAGGCGGCACT 1743

QY 1274 ATGGCTACTCCGTGCGAGCTACTGCGGGGTGGGTGA 1311
Db 1744 ACGGCTACTCCGTGCGAGCTACTGCGGGGTGGGTGA 1781

RESULT 9
LOCUS      AF177906              1480 bp      DNA      linear      BCT 10-SEP-1999
DEFINITION Pyrococcus woesei alpha amylase (amyA) gene, complete cds.
ACCESSION  AF177906
VERSION    AF177906.1  GI:5853153
KEYWORDS   .
SOURCE     Pyrococcus woesei
ORGANISM   Pyrococcus woesei
            Archaea; Euryarchaeota; Thermococci; Thermococcales;
            Thermococcaceae; Pyrococcus.
REFERENCE  1 (bases 1 to 1480)
            Chunlin Lu, Weizheng, J. and Yunyan, Y.
            Cloning and Expression of Alpha Amylase from Pyrococcus woesei
            Unpublished
REFERENCE  2 (bases 1 to 1480)
            Chunlin Lu, Weizheng, J. and Yunyan, Y.
            Direct Submission
            Submitted (15-AUG-1999) New Strategy Biotech Company, 500 Cao Bao
            Road, Shanghai 200233, P.R.China
FEATURES   Location/Qualifiers
            1..1480
                /organism="Pyrococcus woesei"
                /mol_type="genomic DNA"
                /strain="DSM5773"
                /db_xref="taxon:2262"
                98..1480
                    /gene="amyA"
                    98..1480
                        /gene="amyA"
                        /SC_number="3.2.1.1"
                        /note="similar to Pyrococcus furiosus alpha amylase"

```

/codon_start=1 /transl_table=11 /product="alpha amylase" /protein_id="A4D5438.1" /db_xref="GI:5853154" /translation="NNIKKLPTLLTLLFFIVLSPVSAKYLEBEGGVIMQAFYWD VPGGGIWDHIRSKIPVEYAGISAIWLPPSKGMSGGYMDYDFDLGEYQKG TVETRFSGKEELVRLIOAHAYGIKVIADVNNHRRAGDLBNPFGVGYTWDFSKVA SKYTYANLDFHNLHCCDECTFGFPDI CHHKEWDOYWLKSNESVAAVLRISGPD GWRFDYKGVGNVDRDLNHWGWAVEYWDVNDALLSWAYSGAKVDFPFPLYYKM DEAFONNNIPALVYALQNGQTVVSRPFKAVTFVANHTDIIIMNKYPAIFALITYEGQ PVIYDFRFEWLNKDLINLIWHDLAGGSTTIVYDNDDELI FVRNGDSRRPLITY INLSPNWRGVYVYVFKFAGACIHEYTGNLGNWVKRVDSGWSVYLEAPHPDPANGYYG YSVWSYCVGVG"									
Query Match 50.5%; Score 792.8; DB 1; Length 1480; Best Local Similarity 75.4%; Pred. No. 1.3e-168; Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;									
QY	4	GCCAACTACTCCGAGCTCGAAAGGGCGGGGTCAATAATGCAGCGCTTCTACTGGGACGCTG	63						
DB	173	GCAAAATACTTGGAGCTTGAAGGGAGGAGTATAATGCAAGCATTTCTATTGGGATGTT	232						
QY	64	CCTTCAGGAGGAATATGGTGGGACACATACCGCGAAGATACCGGAGTGGTACGATGCC	123						
DB	233	CCAGGGGGAGGAATTTGGTGGGATCATATAAGATCGAAGATTCCTGAATGGTATGAAGCT	292						
QY	124	GGAACTCCGCAATATGGATTCCTCCCGGAGCAAGGGCATGGCGGCGCTTATTCGATG	183						
DB	293	GGAACTCTTGCAATATGGCTACTCCACCAAGCAAGGGATGAGTGGAGGATATTCAATG	352						
QY	184	GGCTACGACCCCTACGACTTCTTTGACCTCGTGGTACGACCAAGGGAACGGTACAG	243						
DB	353	GGCTACGATCCCTATGATTACTTTGATCTCGGCGAGTACTACCAAGGGAACCTGAGAG	412						
QY	244	ACCGGCTTTGGCTCCAGCAGAGCTGCGTAACATGATAAACCAGCCCGCTATCGC	303						
DB	413	ACCGGTTTTGGATCAAAAAGAAAGACTAGTGGATTGATACAACTGCCATGCCCTATGGA	472						
QY	304	ATGAAGGTAATAGCCGATATAGTCATCAACCCGCGCGGGGTGACCTGGAGTGGAAAC	363						
DB	473	ATAAAGGTAATGCCGATGTAGTTATAAACCAAGAGGCTGGTGGTGAACCTGGAAC	532						
QY	364	CCCTTCGTGAACACTATACCTGGACCGACTTCTCAAAGTCCGCTCGGTTAAATACAG	423						
DB	533	CCCTTCGTGGAGATTACACATGGACAGACTTTCTTAAAGTTGCTCAGGGAATATACA	592						
QY	424	GCCAACTACTCGACTTCCACCCGACGAGCTCCATGCGGGGATTCGGGAACATTGGA	483						
DB	593	GCTACTATCTGACTTCCATCCAAACGAGCTTCAITGTTGACGAAGAACCTTTGGA	652						
QY	484	GGCTATCCCGACATATGCCACGACAGAGCTGGACCACTATGGCTCTGGGCCACGACG	543						
DB	653	GGATTTCCAGATATATGTCTATCACAAGAGTGGATCAGTCTGGCTATGGAAGGCAAT	712						
QY	544	GAGAGTACGCGGCATATCTCAGGACATCGGCATCGCTGCGGCTTCGACTACGTC	603						
DB	713	GAGAGTTATGCTGCTTATTTAAGAACATAGATTTGATGGTTGAGATTTCACATGTT	772						
QY	604	AAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGCGGTT	663						
DB	773	AAGGGCTATGGAGCTTGGGTTGTCAGAGACTGGCTTAATTTGGTGGGAGGTTGGGCGTT	832						
QY	664	GAGAGTACTGGGACACCACTCGACGCTGTCTCAACTGGGCATCTCGAGCGGTGCC	723						
DB	833	GAGAGTACTGGGACACCAATGTAGATGCACTACTAAGCTGGGCATATGAGAGTGGTGA	892						
QY	724	AAGGTCTTTGACTTCGCGCTTACTACTAAGAGTGGATGAGCGCTTTTGACAAACAAACATT	783						
DB	893	AAGGTCTTTGACTTCGCGCTTACTACTATAAATGGATGAAGCATTTTGACAAACAAACATT	952						
QY	784	CCAGCGCTGCTCTCTGCGCTTCAGAACGGCCAGACTGTTGTTCTCCGCGACCCGCTCAAG	843						

DB	953	CCAGCATTAAGTCTATGCTATGCTTACAAAACGGACAACTGTAGTTTCGAGAGATCCATTAAAG	1012						
QY	844	GCCGTAACTTTGTAGCAAAACACGACACCGATATAATCTGGAAACAGATATCCAGCCTAC	903						
DB	1013	GCACTAACTTTGCTGCTCCCAATCAATGACAGATATAATATGGAACAGATATCCAGCATAT	1072						
QY	904	GCCTTCATCTCCCTACCTACGAGGGCCAGCCGACAAATATTTCTACCGACTACGAGGAGTGG	963						
DB	1073	GCCTTCATATTGACATATGAGGACAGCCAGTAATATTTCTACAGGACTTTGAGGAATGG	1132						
QY	964	CTCAACAGGATTAAGCTCAAGAACCTTCTCTGATACATGAGAACCTCCGCGGAGGAAGC	1023						
DB	1133	CTGAACAAAGGATTAAGCTTAATTAACCTCATTTGGATCCATGATCATTTGGCAGGAGGAAC	1192						
QY	1024	ACCGACATAGTCTTACTACGATAACGATGAATCTCATCTTCTGTCAGGAACGGCTACGGGGAC	1083						
DB	1193	ACAACAAATTTGCTTACTACGACAAACGATGAGCTCATATTTGTGAGAATGGAGATTCAGA	1252						
QY	1084	AAGCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCCCGAAGTGGGTTTAT	1143						
DB	1253	AGGCTGGGCTTATAACCTTACATTAACCTTGAAGCCCTAACTGGGTTGTTAGTGGGTATAC	1312						
QY	1144	GTCCGGAAGTTCCGCGGGCTGCATCCAGAGTATACCTGTAACCTCGAGGCTGGGTA	1203						
DB	1313	GTTCGGAAGTTCCGCGGGCTTGTATTCATGAATACCTTGGAAACCTAGGAGGATGGGTA	1372						
QY	1204	GACAAAGTACCTTACTCAAGCGGCTGGGCTTATCTCGAAGTCCAGCTTACGACCTGCC	1263						
DB	1373	GATAAAGAGTAGATAGTAGGATGGGTATACCTAGAGCACCACTCAGATCCAGCT	1432						
QY	1264	ACGGGCGATATGGCTACTCCGTGTCGAGCTACTGCGGGGTGGGCTGA	1311						
DB	1433	ACGGGCTACTATGGGTACTCCGTATGGAGTTATTTGGTGGGTGA	1480						
RESULT 10									
AF240464									
LOCUS									
DEFINITION									
ACCESSION									
VERSION									
KEYWORDS									
SOURCE									
ORGANISM									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
FEATURES									
source									
CDS									
1..1561									
/organism="Pyrococcus woesei"									
/mol_type="genomic DNA"									
/db_xref="taxon:2262"									
/notes="DSM 3773"									
137..1519									
/notes="pow"									
/codon_start=1									
/transl_table=11									
/product="alpha-amylase"									
/protein_id="AAF4453.1"									
/db_xref="GI:7271927"									
/translation="NNIKKLPTLLTLLFFIVLSPVSAKYLEBEGGVIMQAFYWD VPGGGIWDHIRSKIPVEYAGISAIWLPPSKGMSGGYMDYDFDLGEYQKG TVETRFSGKEELVRLIOAHAYGIKVIADVNNHRRAGDLBNPFGVGYTWDFSKVA SKYTYANLDFHNLHCCDEGTFGGFPDI CHHKEWDOYWLKSNESVAAVLRISGPD GWRFDYKGVGNVDRDLNHWGWAVEYWDVNDALLSWAYSGAKVDFPFPLYYKM DEAFONNNIPALVYALQNGQTVVSRPFKAVTFVANHTDIIIMNKYPAIFALITYEGQ PVIYDFRFEWLNKDLINLIWHDLAGGSTTIVYDNDDELI FVRNGDSRRPLITY INLSPNWRGVYVYVFKFAGACIHEYTGNLGNWVKRVDSGWSVYLEAPHPDPANGYYG YSVWSYCVGVG"									

SVGRYNYLDFHNLHCCDEGTFGGFPDI CHHKEWDOYWLKSNESVAAVLRISGPD

GWRPDKYGCWVVDLNLNWWGKWAQBEYDNTVDALLSWAYESGAKVDFDPLVYKQ  
DEAFNNNPALVXALQNLVYNDWPPKAVTIVANHTDITWNKYPAYAPILTYEQ  
PVIQFDFEHLNKKLINTWHDHLAGSTTVYVDNDEILFVNGDSRRPGLITY  
ILNSPNWGVYVYPPKACACIHEYTGNLGGWVDKRDSSGVMVLEAPPHDPANGYV  
YSW5V5CVG2"

## ORIGIN

```
Query Match          60.5%; Score 792.8; DB 1; Length 1561;
Best Local Similarity 75.4%; Pred. No. 1.3e-168;
Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;
```

Qy	4	GC	CAAGTACTCGAGCTGGAAAAGGCGGGTCAATAATGCAGCGTTCTACTCGGACGTG	63
Db	212	GC	AAAACTCTGGAGCTTGAAGAGGGAGGATTAATAATGCAAGCATTTCTATTGGGATGTT	271
Qy	64	CC	TTACAGAGCAATATGGTGGGACACAATAACGGCAGAAAGTACCGGAGTGGTACGATGCC	123
Db	272	CC	ACGGGAGGAATTTGGTGGGATCATATAAGATCGAAGATTCCTGAATGGTATGAAGCT	331
Qy	124	GG	AATCTCCGCAATATGGAATTCGCCCGGCGAGCAAGGSCATGGCGCGCGCTATTGCGATG	183
Db	332	GG	AATCTCTGCAATATGGCTACCTCCAACCAAGCAAGGGGATGATGGAGGATATTCAATG	391
Qy	184	GG	CTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGAAACGGTAGAG	243
Db	392	GG	CTACGATCCCTATGATTACTTTTGATCTCGGCGAGTACTTCCAGAGGGAACTGTAGAG	451
Qy	244	AC	CGCTTTGGCTCCAAAGCAGGAGCTCGTGAAATGATAAACACCGCCACGCTATGGC	303
Db	452	AC	CGTTTTGGATCAAAAGAAAGAACTAGTGAGATTGTATACAAATGCCCCATGCTATGGA	511
Qy	304	AT	GAGGTAAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTCGGAGTGGAAAC	363
Db	512	AT	AAGGTAAATGCCGATGTAGTTATAAACCAACAGGGCTGGTGGTGAACCTTAGAATGGAAC	571
Qy	364	CC	TTTCGTGAACGACTATACCTGSCACCGACTTCTCAAAGTTCGGTCCGGTAAATAACAG	423
Db	572	CC	TTTCGTGGAGATTACATGACACAGACTTTTCTAAAGTTGCTCAGGCAATAATACA	631
Qy	424	GC	CAACTACTCTGACTTTCACCCGAAACGAGCTCCATCGGCGGATTCGGAAACATTTGGA	483
Db	632	GC	TAACTACTGTGACTTTCATCCAAACGAGCTTCATTGTTGTGACGAAGGAACCTTTGGA	691
Qy	484	GG	CTATCCGACATATCCACGACCAAGAGCTGGGACCAAGTACTGGCTCTGGGCCAGCAG	543
Db	692	GG	ATTTCCAGATAATGTCTATCAAAAGAGTGGGATCAGTACTGGCTATGGAAGAGCAAT	751
Qy	544	GAG	AGCTACGGGCAATATCTCAGAGAGCATCGGCATCGATCGCTGGCGCTTCGACTAGCTC	603
Db	752	GAG	AGTTATGCTGCTTATTTAAGAAGCATAGGATTTGATGGTTGGAGATTGCACTATGTT	811
Qy	604	AAG	GGCTATGCTCCCTGGGTGCTCAAGGAGCTGGCTGAACCTGGTGGGAGAGCTGGCGGTT	663
Db	812	AAG	GGCTATGAGCTTGGGTTGTCAGAGACTGGCTTAATGGTGGGAGGTTGGGCAGTT	871
Qy	664	GG	AGTACTGGGACACAACGTCGACGCTGTTCTCAACTGGGCACTACGAGCGGTGCC	723
Db	872	GG	AGTACTGGGACACAATAATGATAGTGCATCTAATAAGCTGGGCATATGAGAGTGTGCA	931
Qy	724	AAG	GTCTTTGACTTTCGCCCTCTACTACAAAGATGGATGAGGCTTTGACAAACAAAACATT	783
Db	932	AAG	GTCTTTGACTTTCGCCCTCTACTATAAATGATGAAGCAATTGACATTAACAACATT	991
Qy	784	CC	AGCGCTCGTCTCTGCCCTTCGAAACGGCGCAGACTGTTGTTCTCCCGGACCCGTTCAAG	843
Db	992	CC	AGCATTAGTCTATGCGCTTACAAAACGGGCAAACTGTAGTTTCGAGAGATCCCAATTAAG	1051
Qy	844	GC	GTAACTTTGTAGCAAAACCAAGCACCGATATAATCTGGAAACAAGTATCCAGCGCTAC	903
Db	1052	GC	AGTAACTTTGTTGCCAATCATGACACAGATATAATATGGAACAAGTATCCAGCATAT	1111
Qy	904	CG	GTTCATCTCACCTACGAGGGCCAGCGCAATAATTTCTACCGGACTACGAGGAGTGG	963

Db	1112	GCGTTCATATTGACATATAGGGGACCCGAGTAATATTCTACAGGCATTTTAGGAATGG	1171
Qy	964	CTCAACAAGGATAAGCTCAAGAACCCTCATCTGGATACATGAGAACCTCGCCGAGGAAGC	1023
Db	1172	CTGAAACAAGATTAAGCTAATTACCTCATTTGGATCCATGATCATTTGGCAGAGGAAGC	1231
Qy	1024	ACCGACATAGTCTACTACGATAACGATGAACATCTTTCGTACAGGAACGGCTACGGGGAC	1083
Db	1232	ACAACAATTGTCTACTACGACAACGATGAGCTCATATTTGTGAGAAATGGAGATTTCTAGA	1291
Qy	1084	AAGCCGGGGCTTATAACCTTACATCAACCTTAGGCTCGAGCAAGGCCGGAAGGTGGGTTTAT	1143
Db	1292	AGCCCTGGGGCTTATAACTTACATTAACCTTGAGCCCTAACCTGGGTTGGTAGTGGGTATAC	1351
Qy	1144	GTCCGGAAGTTCCGGGGCGGTGCATCCAGAGTATACTGGTAACCTCGAGAGCTGGGTA	1203
Db	1352	GTTCGAAAGTTTCAGGGGCTTGTTNTTCATGAATACCTGGAAAACCTAGGAGGATGGGTA	1411
Qy	1204	GACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCCTGCC	1263
Db	1412	GATAAAGAGTAGATAGTAGCGGATGGGTATACCTAGAGSCACCACTCACGATCCAGCT	1471
Qy	1264	AACGGGAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGGTGA	1311
Db	1472	AACGGGTACTATUGGGTACTCCGPATGGAGTTATTGGTGTTCGGTGCA	1519

## RESULT 11

PFU95622	PFU95622	1740 bp	DNA	linear	BCT 27-AUG-1997
LOCUS					
DEFINITION	Pyrococcus furiosus	alpha-amylase gene,		complete cds.	

ACCESSION	U96622	GI:2343082
VERSION	U96622.1	

**KEYWORDS**  
SOURCE  
ORGANISM  
Pyrococcus furiosus  
Pyrococcus furiosus  
Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; Pyrococcus.

REFERENCE  
1 (bases 1 to 1740)  
AUTHORS  
Jorgensen, S., Vorglas, C.E. and Antranikian, G.  
TITLE  
Cloning, sequencing, characterization, and expression of an  
extracellular alpha-amylase from the hyperthermophilic archaeon  
*Pyrococcus furiosus* in *Escherichia coli* and *Bacillus subtilis*  
JOURNAL  
J. Biol. Chem. 272 (26), 16335-16342 (1997)

**MEDLINE** 97341170

PUBMED 9195939

REFERENCE 2 (bases 1 to 1740)

**AUTHORS**  
Joergensen, S.T.,  
Director, Publications

**JOURNAL TITLE**  
Submitted (07-APR-1997) Bacterial Gene Technology, Enzyme Research,  
Direct Submission  
Novo Nordisk A/S, Novo Alle, Bagsvaerd 2880, Denmark

## FEATURES

```

source
1. .1740
/organism="Pyrococcus furiosus"
/mol_type="genomic DNA"
/strain="DSM3638"
/db_xref="taxon:2261"
229. .1611
CDS

```

505

```

/codon_start=1
/transl_table=11
/product="alpha-amylase"
/protein_id="AAB67705.1"
/db_xref="GI:2343083"
/translation="MNKIKLPILTLTLLFFIVLSPVSAAKYLEELEGGVIMQAFYWD
VPGGIGIWHIRSKIPETWEAGISAIWLPPSGKMGSGYSGMDYDPYDFDLGSGYQKG
VSTRYGSXEELVRLICTAHAYGIVKIVAVINHRAGGDLEWNPFFVDYDTWTDFSKVA
SGYTYALNDFHFNELHCCDEGTFGGFPDICHKENDQYWLKSNESYAAYLESIGFI
GWRFDYKYGAWVWDNLNWNWGVAWGYNVDNLALSWSAYESAUKVDFDFPLYKIM
DEAFDNNNIPALVYALQNGTVPDRDPFKAVTFVANHDTIIWNKYTAPAFILTYEGQ
PVTFYDFPEWLNKKLIINLWTHDFLAGGSTIIVTDNDLIFVANGDSRRPGLTY
INLSPNNGRWGVVPKFPAGACIHEYTCNLGGWWDKRVDSGSGWYLEAPPHDPANGYYG
YSVSYCYGVG"
```

## ORIGIN



Query Match 60.5%; Score 792.8; DB 1; Length 1740;  
 Best Local Similarity 75.4%; Pred. No. 1.3e-168;  
 Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 4 GCCAAGTACTCCGAGCTGGAAGGCGGGGTCATATCGAGCGTTCCTACTGGGACGTG 63  
 DB 304 GCAAAATATCTGGAGCTTGAAGGAGGAGGATTAATGCAAGCAATCTTATTTGGGATGTT 363  
 QY 64 CTTTCAGGAGGATATGTTGGGAGGACATACCGGAGAGATACCGGAGTGTGACATGCC 123  
 DB 364 CCAGGGGAGGAGATTTGGTGGGATCATATAAGATCGAAGATTCCTGAATGGTATGAAGCT 423  
 QY 124 GGAATCTCCGCAATATGGAATCCCGCGGACCAAGGCGATGGCGCGCTATTCGATG 183  
 DB 424 GGAATCTCTGCAATATGGCTACTCCACCAAGAGGGATGAGTGGAGGATATTCATG 483  
 QY 184 GGTACGACCCCTACGACTTCTTGAACCTCGGTGAGTACGACCAAGAGGAAAGGTAGAG 243  
 DB 484 GGTACGATCCCTATGATTAATCTCGGAGTACTACCAAGAGGAACTAGAG 543  
 QY 244 ACGCGTTTGGCTCCAGCAGAGCTCGTGAACATGAATAACACCGCCAGCCTATGCC 303  
 DB 544 ACGCGTTTGGATCAAGAGAACTAGTGAATGATCAAACTGCCATGCCCTATGGA 603  
 QY 304 ATGAAGGTAATAGCCGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGAAC 363  
 DB 604 ATGAAGGTAATCGCGATAGTATTAACACAGCGGTGGTGTGACCTAGATGGAAC 663  
 QY 364 CCTTTCGTAACGACTATACCTGGACCGACTTCTCAAGGTCGCTCGGTGAATACAG 423  
 DB 664 CCTTTCGTTGGAGATTACACATGGAAGCACTTTCTAAAGTTGCTCAGGGAATATACA 723  
 QY 424 GCCAATCTACTCGACTTCCACCGCAACAGCTCCATCGCGGCGATTCGGGAACATTTGGA 483  
 DB 724 GCTAATATCTGGACTTCCATCCAAACAGCTTCATTTGTGAGCAAGGAACCTTTGA 783  
 QY 484 GGTATCCGCAATATGCGACAGACAGAGCTGGGACAGTACTGCTCTGGGCGAGCCAG 543  
 DB 784 GGAATTTCCAGATATATGTATCAAAAGATGGGATCAGTACTGGCTATGGAAGCAAT 843  
 QY 544 GAGAGCTACGCGCATATCTCAGGAGCATCGGCATCGATCGCTGGCTTCGACTACGTC 603  
 DB 844 GAGAGTATGCTGCTTATTTAAGAGCATAGATTTGATGTTGAGATTTGACTATGTT 903  
 QY 604 AAGGCTATGCTCCCTGGGTCGTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGGTT 663  
 DB 904 AAGGCTATGGAGCTTGGGTTGTCAAGAGCTGGCTTAATTTGGTGGGAGGTTGGCGAGTT 963  
 QY 664 GGAGGATCTGGGACACCAACCTCGACGCTGTTCTCACTGGGATCTCGAGCGGTGCC 723  
 DB 964 GGAGGATCTGGGACACCAATGTAGATGCACCTACTAAGCTGGGATATAGAGTGGTGCA 1023  
 QY 724 AAGGCTTTTGACTTCGCCCTCTACTACAGATGATGAGGCTTTTGACAAACAAACATT 783  
 DB 1024 AAGGCTTTTGACTTCGCCCTCTACTATAAATGATGAGCAATTTGACAAATACCAATT 1083  
 QY 784 CAGCGCTGCTCTGCCCTTCAGAACCGCAGACTGTTGTCCTCCGCGGACCCGTTCAAG 843  
 DB 1084 CCAGCAATTAAGTATGATGCTCAAAACCGCAAACTGTAGTTTCAGAGATTCATTTAAG 1143  
 QY 844 GCGGTAACTTTGTAGCAACACGACACCGATATAATCTGGAAACAGATATCCAGCTTAC 903  
 DB 1144 GCAGTAACCTTTGTTGCCATCATGACACAGATATAATATGAAACAGATATCCAGCAT 1203  
 QY 904 GGTTCATCTCCTACCTACGAGGCGCAGCGCAATATTTCTACCGGCACTACGAGGATGG 963  
 DB 1204 GCGTTTATATGACATATAGGAGGACAGCCAGTAATATTTCTACAGGCACTTTGAGGAATGG 1263  
 QY 964 CTCACAGGATTAAGCTCAAGAACTCATCTGGATACATGAACTCGCGGAGGAGGAGC 1023  
 DB 1264 CTGACAAAGGATTAAGTATTAACCTCATTTTGGATCCATGATCATTTTGGCAGGAGGAGC 1323

QY 1024 ACCGACATAGTCTACTACGATAACGATGAACCTCATCTTCGTACGAAACGGCTACGGGAC 1083  
 DB 1324 ACAACATATGTTCTACTACGACACGATGAGCTCATATTTGTGAAATCGAGATTCTAGA 1383  
 QY 1084 AAGCCGGGGTATTAACTACATCAACCTAGGCTCGAGCAAGCCGGAAGGTGGGTTTAT 1143  
 DB 1384 AGGCTGGGCTTATACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1443  
 QY 1144 GTCCGAGTTCGGGCGGCTGCATCCAGAGTATACCTGTAACCTCGAGGCTGGGTA 1203  
 DB 1444 GTTCCAAAGTTTCGAGGGGCTTGTATTCATGAATACACTGGAACCTAGAGATGGGTA 1503  
 QY 1204 GACAAAGTACCTCTACTCAAGCGCTGGGCTTATCTCGAAGCTCCAGCTTACGACCCCTGCC 1263  
 DB 1504 GATAAAAGAGTAGATAGTAGCGATGGTATACCTAGAGCACCACTCACGATCCAGCT 1563  
 QY 1264 AAGCGGAGTATGGCTACTCCGCTGGAGTACTGCGGGGTGGGCTGA 1311  
 DB 1564 AACGGTACTATGGTACTCCGCTATGGAGTATTTGGTGGTGA 1611

RESULT 12  
 E09372  
 LOCUS E09372 2134 bp DNA linear PAT 29-SEP-1997  
 DEFINITION Heat-resistant alpha-amylase gene.  
 ACCESSION E09372  
 VERSION E09372.1 GI:22025999  
 KEYWORDS JP 1995143880-A/1.  
 SOURCE Pyrococcus furiosus  
 ORGANISM Pyrococcus furiosus  
 Archaea; Euryarchaeota; Thermococci; Thermococcales;  
 Thermococcaceae; Pyrococcus.  
 REFERENCE 1 (bases 1 to 2134)  
 AUTHORS Taguchi,Y., Nishikawa,M., Koyama,N., Yamamoto,K., Asada,K. and Katou,I.  
 TITLE ALPHA-AMYLASE GENE HAVING ULTRA-HIGH HEAT-RESISTANCE  
 JOURNAL Patent: JP 1995143880-A 1 06-JUN-1995;  
 TAKARA SHUZO CO LTD  
 COMMENT OS Pyrococcus furiosus  
 PN JP 1995143880-A/1  
 PD 06-JUN-1995  
 PF 18-AUG-1994 JP 1994215253  
 PR 01-OCT-1993 JP 93P 267777  
 PI TAGUCHI YUKI, NISHIKAWA MIO, KOYAMA NOBUHIRO, PI YAMAMOTO KATSUHIKO,  
 ASADA KIYOZOU, KATOU IKUNOSHIN  
 PC C12N15/09,C12N9/28,C12N15/09,C12R1:01,(C12N9/28,C12R1:19);  
 CC strandedness: Double;  
 CC topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1. .2134 /organism='Pyrococcus furiosus' FT promoter  
 FT RBS 337. .373 380. .388  
 FT CDS 395. .1777 /product='heat-resistant alpha-amylase FT  
 FT precursor,  
 FT sig\_peptide 395. .469  
 FT mat\_peptide 470. .1774 /product='heat-resistant alpha-amylase'.  
 FT  
 FEATURES  
 source 1. .2134 Location/Qualifiers  
 /organism='Pyrococcus furiosus'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:2261'  
 ORIGIN  
 Query Match 60.5%; Score 792.8; DB 6; Length 2134;  
 Best Local Similarity 75.4%; Pred. No. 1.3e-168;  
 Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;  
 QY 4 GCCAAGTACTCCGAGCTGGAAGGCGGGGTCATATCGAGCGTTCCTACTGGGACGTG 63



Db 470 GCAAAATACCTTGGAGCTTGAAGAGGAGGAGTTAATGCAAGCAATCTAATGGGATGTT 529  
QY 64 CTTTCAGAGGAATATGTTGGGACAAATACGGCAGAGATACCGGAGTGTACGATGCC 123  
Db 530 CCAGGGGAGGAATTTGGTGGGATCATATAGATCGAAGATTCCTGAATGTATGAAGCT 589  
QY 124 GGAATCTCCGCAATATGATTTCCCGCGGAGCAGGAGGATGGCGGCGCTATTTCGATG 183  
Db 590 GGAATCTCTGCAATATGCTACCTCCACCAAGCAGGGGATGAGTGGAGGATATTCAATG 649  
QY 184 GGTACGACCCCTACGACTTTTTCACCTCGGTGATGACGACCAAGAGGAAACGGTAGAG 243  
Db 650 GGTACGATCCCTATGATTTTCTGCGGAGTACTACCAAGAGGAACTGTAGAG 709  
QY 244 AGCGCTTTGGCTCAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTATGGC 303  
Db 710 ACGCGTTTGGATCAAAAGAGNACTAGTGAATGATACAACTGCCCTATGGA 769  
QY 304 ATGAAGGTAATAGCCGATATAGTCAACACCGCGCGGCGGTGACCTGGAATGGAAC 363  
Db 770 ATAAAGGTAATCGCGATGTATGTTATAAACCACAGCGGCTGGTGTGACCTAGAATGGAAC 829  
QY 364 CCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTCGCTCGGTAATAATACG 423  
Db 830 CCCTTCGTGGAGATTACATGACAGACTTCTAAAGTTGCTCAGGGAATATACA 889  
QY 424 GCCAACTACCTGGACTTCCACCCGACGAGCTCCATCGCGGCGATTCGGAACATTTGGA 483  
Db 890 GCTAACTATCTGGACTTCCATCCAAACGAGCTTCATTTGTGACGAAGGAACCTTTGGA 949  
QY 484 GGTATCCGACATATGCGACGACGAGCTGGGACGAGTACTGGCTCGGCCACGACGAG 543  
Db 950 GGAATTCAGATATATGTCATCAACAAAGATGGGATCAGTACTGGCTATGGAAGAGCAAT 1009  
QY 544 GAGACTACGGGGATATCTCAGAGACTCGGATCGGATCGCTGGCGTTCGACTACGTC 603  
Db 1010 GAGATTTATGCTGCTATTTAAGACATAGATTTGATGTTGGAGATTGACTATGTT 1069  
QY 604 AAGGCTATGCTCCCTGGGTGCTCAAGACTGGGTGAACCTGGTGGGAGGCTGGCGGGT 663  
Db 1070 AAGGCTATGAGCTTGGGTTGTGAGAGACTGGCTTAATTTGGTGGGAGGTTGGGCGATT 1129  
QY 664 GGAGAGTACTGGGACCAACGTCGACGCTGTTCTCAACTGGGATACTCGAGCGGTGCC 723  
Db 1130 GGAGAGTACTGGGACCAAAATGTAGATGCACTACTAAGCTGGGCATATGAGAGTGTGCA 1189  
QY 724 AAGGCTTTGACTTCGCCCTCTACTACAAGATGGATAGGCGCTTTGACAAACAAACATT 783  
Db 1190 AAGGCTTTGACTTCGCCCTCTACTATAAATGGATGAAGCATTTGACAAATACACATT 1249  
QY 784 CCAGCGCTCGTCTCGCTTCAGAACGCGCAGACTGTTGTTCTCCCGGACCGCTTCAAG 843  
Db 1250 CCAGCATATGCTATGCGCCCTACAAACGGAACAACTGTAGTTTCGAGAGATCCATTAAAG 1309  
QY 844 GCGTAACCTTTGTAGCAAAACACGACCGATATAATCTGGAACAAGTATCCAGCGCTAC 903  
Db 1310 GCAGTAACCTTTGTTGCCATCATGACACAGATATAATATGAACAGTATCCAGCATAT 1369  
QY 904 GGTTCATCTCTACCTAGAGGCGAGCGGACATATTTCTACCGGACTACGAGAGTGG 963  
Db 1370 GCGTTTCAATTTGACATATGAGGACGAGCAGTAATATTTCTACAGGGACTTTGAGGAATGG 1429  
QY 964 CTCACAGAGGTAAGCTCAAGAACTCATCTGGATACATGAGAACCTCGCGGAGGAGGAGC 1023  
Db 1430 CTGAACAGAGTAAGCTAATTAACCTCATTTGGATCCATGATCATTTGGCAGAGGAGGAC 1489  
QY 1024 ACCGACATAGTCTACTACGATAACGATGAATCATCTTTCTGAGAAACGGCTACGGGAC 1083  
Db 1490 ACAACATTTGCTACTAGCAACAGATGAGCTCATATTTGTGAGAAATGGAGATTCTAGA 1549  
QY 1084 AAGCGGGGCTTATTAACCTATACCTAGCTCGACCAAGCGCGGAGGTGGTTTAT 1143

Db 1550 AGCCCTGGGCTTATACTTACATTAACCTTGAAGCCCTAACTGGGTTCGTAGTGGGTATAC 1609  
QY 1144 GTCGCCAGGTTCCGCGCGCGTGCATCCAGAGTACTGCTAACTCCGAGGCTCGGTA 1203  
Db 1610 GTTCCAAAGTTTCGACGGGCTTGTATTCATGATACACCTGGAACCTAGAGGATGGTA 1669  
QY 1204 GACAAGTACGCTTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCCCTGCC 1263  
Db 1670 GATAAAGAGTAGATAGTAGCGGATGGGTATACCTAGAGGACCACTCAGGATCAGCT 1729  
QY 1264 AAGCGGCTATGGCTACTCCGCTGTGAGCTACTCGGGGTGGGCTGA 1311  
Db 1730 AACGGCTACTATGGGTACTCCGCTATGGGTATTTGGGTGGA 1777

## RESULT 13

AF001268

LOCUS

Pyrococcus furiosus 2627 bp DNA linear BCT 24-SEP-1997

ACCESSION AF001268

VERSION AF001268.1 GI:2183105

KEYWORDS

SOURCE

ORGANISM

Pyrococcus furiosus

Archaea; Euryarchaeota; Thermococci; Thermococcales;

Thermococcaceae; Pyrococcus.

REFERENCE 1 (bases 1 to 2627)

Dong, G., Vieille, C., Savchenko, A. and Zeikus, J. G.

Cloning, sequencing, and expression of the gene encoding

extracellular alpha-amylase from Pyrococcus furiosus and

biochemical characterization of the recombinant enzyme

Appl. Environ. Microbiol. 63 (9), 3569-3576 (1997)

JOURNAL

MEDLINE

PubMed

REFERENCE 2 (bases 1 to 2627)

Dong, G., Vieille, C., Savchenko, A. and Zeikus, J. G.

Direct Submission

Submitted (24-APR-1997) Biochemistry, Michigan State University,

410 Biochemistry Building, East Lansing, MI 48824, USA

JOURNAL

FEATURES

source

1. 2627

/organism="Pyrococcus furiosus"

/mol\_type="genomic DNA"

/strain="DSM 3638"

/db\_xref="taxon:2261"

676..2058

/gene="amyA"

/EC\_number="3.2.1.1"

/note="alpha-1,4-glucan 4-glucanohydrolase; acidic

extracellular enzyme"

/codon\_start=1

/transl\_table=11

/product="alpha amylase"

/protein\_id="AAC45663.1"

/db\_xref="GI:2183106"

/translation="MNKKLTLLTLLFFVLASPVSAKYLEEGGVINQAFVND

VPGGLWDRHRSKIPETAGISALWLPSPKSGSGSMGIDPDYDFDLGEYQKG

TVEFTRGSEBELVRIQTAYGKIADVINHRAGGLEWNPVFGDVTWDFSKVA

SKQYATYLDLPHNELCCDGTGGFDICHHKENDQYWLKSNESYAAIYLRSGFA

GWRFDYKYGAVVVDMLNMWGWAGVYDNTDALISWAYESGAKVDFPLPYKM

DFAFDNNIPALVALQNGQTVVSPKAVTFVANHDTDIWNKYPAVAFILTYEGQ

PIVFYDFREELNKKLINLIWIHDLAGGSTTVYDNDLIVRNGSDSRPGLTY

INLSNFWGRVYVPKFPAGACIHEVTGNLGNVDRVDSGHWYLEAPHPDPANGYIG

ISVWSYCGVG"

## ORIGIN

Query Match 60.5%; Score 792.8; DB 1; Length 2627;  
Best Local Similarity 75.4%; Pred. No. 1.4e-168;  
Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 4 GCCAAGTACTCCGAGCTGGAAAGGCGGGGCTCATTAATGACGCGTTCTTACTGGGACGTG 63

751 GCAAAATACCTTGAGCTTGAAGGGAGGATTAATAATGCAAGCATTTCTATTGGGATGTT 810  
 64 CTTTACGAGGAAATATGGTGGGACACATACGGCAGAGATACCGAGTGGTACGATGCC 123  
 811 CCAGGGGAGGAATTTGGTGGGATCATATAAAGATCGAAGATTCCTGAAATGGTATGAAGCT 870  
 124 GGAATCTCCGCAATATGGATTCCTCCCGGAGCAAGGGGATCGGGGGGCGCTATTTCGATG 183  
 871 GGAATCTCTGCAATATGGCTACTCCACCAAGCAGGGGATGAGTGGAGGATATTCAATG 930  
 184 GGTACGACCCCTACGATCTCTTTGACCTCGTGGTACGACAGAGGGAACGGTAGAG 243  
 931 GGTACGATCCCTATGATTTGATCTCTCGGAGTACTACCAAGGGAACCTGTAGAG 990  
 244 AGCGCTTTGGCTCCAGCAGAGCTCGTGAACATGAATAACACCGCCACGCTTATGCG 303  
 991 AGCGCTTTGGATCAAAAGAGAACTAGTGAATGATACAACTGCCATGCCCTATGGA 1050  
 304 ATGAAGGTAAATAGCCGATATAGTCATCAACCCGCGCGCGGTGACCTGGAGTGGAAAC 363  
 1051 ATAAAGGTAAATGCCGATGTAGTTATAAACCAACAGGGCTGGTGGTGAACCTAGATGGAAC 1110  
 364 CCCTTCGTGAACATATACCTGGACCGACTTCTCAAGGTGCGTCCGGTAAATACAG 423  
 1111 CCCTTCGTGGAGATTACATGACAGACTTTTCTAAAGTTGCCCTCAGGGAATATACA 1170  
 424 GCCAATACCTGACCTTCCACCCGAGAGCTCCATGCGGGGATTCGCGAACTTTGGA 483  
 1171 GCTAACTATCTGACTTCCATCCAAAGAGCTTCATTGTTGACGAAGAACTTTGGA 1230  
 484 GGTATCCCGACATATGTCACGACAGAGCTGGGACCATGACTTGGCTCTGGCCGACGAG 543  
 1231 GGAATCCAGATATATGTCATCAAAAGAGTGGATCAGTCTGGCTATGGAAGAGCAAT 1290  
 544 GAGAGCTACCGGCATATCTCAGAGCATCGGCATCGATCGCTGGCGCTTCGACTACGTC 603  
 1291 GAGAGTTATGCTCTATTATTAAGACATAGGATTTGATGGTTGAGATTGCACTATGTT 1350  
 604 AAGGGCTATGCTCCCTGGTGTCAAGGACTGGCTGAACTGGTGGGAGCGTGGCGGTT 663  
 1351 AAGGGCTATGGAGCTTGGGTTGTCAAGACTGGCTTAAATGGTGGGAGGTTGGCGAGTT 1410  
 664 GAGAGTACTGGGACACCAACGACGCTGCTCTCAACTGGGCTACTCGAGCGGTGCC 723  
 1411 GAGAGTACTGGGACACCAATGATGATGACTACTAAGCTGGGCTATGAGAGTGGTGA 1470  
 724 AAGGTCTTTGACTTCGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAACATTT 783  
 1471 AAGGTCTTTGACTTCGCTCTACTATAAATGGATGAAGCATTTGACAAATAACCAATTT 1530  
 784 CAGCGCTGCTCTGCTGCTTGAAGCGGACGAGCTGTTCTCCCGACCGCTTCAAG 843  
 1531 CAGCATTAAGTCTATGCTTACAAACCGGACAACTGTAGTTTCGAGAGATCCATTTAAG 1590  
 844 GCGTAACCTTTGTAGCAAAACCAACGACCGATATATCTGGAACAAAGTATCCAGCTAC 903  
 1591 GCGTAACCTTTGTTGCCATCATGACACAGATATATAATGGAACAGTATCCAGCATAT 1650  
 904 GCGTTCATCTCTACCTACGAGGCGCAGCGACAAATATCTACGCGACTACGAGGATGG 963  
 1651 GCGTTCATATGACATATGAGGACAGCCAGTAAATTTCTACAGGACTTTGAGGAATGG 1710  
 964 CTCACAGAGGATAGCTCAAGAACCTCATCTGATATACATGAGAACTCCCGGAGGAAGC 1023  
 1711 CTGAACAGAGTAAAGCTAATTAACCTCATTTGATCCATGATCATTTGACAGAGGAAGC 1770  
 1024 ACCGACATAGTCTACTACGATACGATGAACCTCATCTTCTGTCAGGAACGGCTACGGGAC 1083  
 1771 ACACAATTTGCTACTACGACCAAGTACGCTCATATTTGTCAGAAATGGAGATTTCTAGA 1830  
 1084 AAGCGGGCTTATACCTACATCACTACCTAGGCTCGAGCAAGCGCCGAAGGTTGGTTAT 1143

1831 AGGCCTGGGCTTATACTTACATTAACCTTGAAGCCCTAAGCTGGTGGTGGGTATATAC 1890  
 1144 GTGCCGAAGTTCCGGGGCGGTGCATCCAGAGTATACTGGTAACCTCGAGGCTGGTA 1203  
 1891 GTTCCAAAGTTTCAGGGGCTTGTATTATCATGATACACTTGGAAACCTAGGAGATGGTA 1950  
 1204 GACAAGTACCTCTACTCAAGCGGCTGGTGTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
 1951 GATAAAAGATAGATAGTAGTGGATGGTATACCTAGAGCACCACCTCAGCATCCAGCT 2010  
 1264 AACGGGAGTATGGCTACTCCGTGTGGAGCTACTGCGGGGTGGGCTGA 1311  
 2011 AACGGCTACTATGGTACTCCGTATGAGTATTGTTGGTGGTGA 2058  
 RESULT 14  
 BD012151/c  
 LOCUS BD012151 3519 bp DNA linear PAT 02-AUG-2002  
 DEFINITION Polypeptide.  
 ACCESSION BD012151  
 VERSION BD012151.1 GI:22092340  
 KEYWORDS WO 0109348-A/4.  
 SOURCE Pyrococcus furiosus  
 ORGANISM Pyrococcus furiosus  
 Archaea; Euryarchaeota; Thermococci; Thermococcales;  
 Thermococcaceae; Pyrococcus.  
 REFERENCE 1 (bases 1 to 3519)  
 AUTHORS Koyama,N., Okui,T., Takakura,H., Asada,K. and Kato,I.  
 TITLE Polypeptide  
 JOURNAL Patent: WO 0109348-A 4 08-FEB-2001;  
 TAKARA SHUZO CO LTD,NOBUTO KOYAMA, TOSHITAKE OKUI, HIKARU TAKAKURA,  
 KIYOZO ASADA, IKUNOSHIN KATO  
 COMMENT OS Pyrococcus furiosus  
 PN WO 0109348-A/4  
 PD 08-FEB-2001  
 PF 26-JUL-2000 WO 2000JP004956  
 PR 02-AUG-1999 JP 99P 218778  
 PI NOBUTO KOYAMA, TOSHITAKE OKUI, HIKARU TAKAKURA, KIYOZO ASADA, PI  
 IKUNOSHIN KATO  
 PC C12N15/56, C12N9/26, C12P19/14, C12N1/15, C12N1/19, C12N1/21, PC  
 C12N5/00  
 CC  
 FH Key Location/Qualifiers.  
 Location/Qualifiers  
 1..3519  
 /organism="Pyrococcus furiosus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:2261"  
 FEATURES  
 source  
 ORIGIN  
 Query Match 60.5%; Score 792.8; DB 6; Length 3519;  
 Best Local Similarity 75.4%; Pred. No. 1.4e-168;  
 Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;  
 4 GCCAAGTACTCCGAGCTGGAAAAGGGCGGGGTCTAATGCGGGCTTCTACTGGGAGTG 63  
 3427 GCAAAATACCTTGGAGCTTGAAGAGGAGGATTAATGCAAGCATTTCTATTGGGATGT 3368  
 64 CTTTCAGAGGATATCGTGGGACACATACGGCAGAGATACCGAGTGGTACGATGCC 123  
 3367 CCAGGGGAGGAAATTTGGTGGGATCATATAAGATCGAAGATTTCTGATGGTATGAGCT 3308  
 124 GGAATCTCCGCAATATGGATTTCCCGCGGAGCAAGGGCATGGCGGCGCTATTTCGATG 183  
 3307 GGAATCTCTGCAATATGGCTACTCCACCAAGCAAGGGGATGAGTGGAGGATATTCAATG 3248  
 184 GGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGGAACGGTAGAG 243  
 3247 GGCTACGATCCCTATGATTTACTTTGATCTCGGCGAGTACTACCAAGAGGAACTGTAGAG 3188  
 244 ACGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACCCGCCACGCTATGCC 303  
 3187 ACGGCTTTGGATCAAAAGAGAACTAGTGAGATTGATACAAACTGCCCATGCTATGGA 3128

QY	304	ATGAGGTAATAGCCGATATAGTATCAACACCAACCGCGCGGTGACCTGGAGTGAAC	363
Db	3127	ATAAAGTAATCGCGATGTAGTTATAAACACACAGCGCTGTGTGACCTAGATGAAC	3068
QY	364	CCCTTCGTGAACGACTATACCTGACCGACTTCCAAAGGTCCGTCGGTAAATACACG	423
Db	3067	CCCTTCGTGGAGATACACATGGACAGACTTTCTAAAGTTGCCCTCAGGGAATATACA	3008
QY	424	GCCAACTACCTCGACTTCCACCCGACGAGCTCCATCGGGCGATTCGCGAACATTTGGA	483
Db	3007	GCTAACTATCTGGACTTCCATCCAAACGAGCTTCATTGTTGACGAAGAACCTTTGGA	2948
QY	484	GCCTATCCGACATATGCCACGACGAGCTGGACAGTACTGGCTCTGGGCCAGCCAG	543
Db	2947	GGATTTCAGATATATGATCATCACAAAGAGTGGGATCAGTACTGGCTATGGAAGACCAAT	2888
QY	544	GAGACTACGGCGCATATCTCAGGAGCATCGGCATCGATGCTGGCGTTCGACTACGTC	603
Db	2887	GAGAGTATGCTGCTTATTAAAGACATAGAAATTGATGTTGGAGATTGACTATGTT	2828
QY	604	AAGGCTATGCTCCCTCGGTGCTGAAGACTCGCTGAACCTGGTGGGAGGCTGGCGGTT	663
Db	2827	AAGGCTATGGAGCTTGGGTTGTACAGACTGGCTTAATTGGTGGGAGGTTGGGCAGTT	2768
QY	664	GGAGAGTACTGGACACCAACGTCGACGCTGTCTCAACTGGGCATCTCGAGCGGTGCC	723
Db	2767	GGAGAGTACTGGACACCAAAATGTAGATGCACTACTAAGCTGGGCATATGAGAGTGTGCA	2708
QY	724	AAGGCTTTGACTTCGCCCTCTACTACAAGATGGATGAGGCTTTGACAAACAAACATT	783
Db	2707	AAGGCTTTGACTTCGCCCTCTACTATAAATGGATGAACATTGACAAATAACACATT	2648
QY	784	CCAGCGCTCGTCTGCGCTTCAGAACCGCCAGACTGTCTCCCGGACCGGTTCAAG	843
Db	2647	CCAGCATTAGTCTATGCCCTCAAAACGGCAAACTGTAGTTTCGAGAGATCCATTTAAG	2588
QY	844	GCCGTAACCTTTGAGCAAAACACACACCGCATATAATCTGGAAACAGATATCCAGCCTAC	903
Db	2587	GCAGTAACCTTTGTTGCCAATCATGACACAGATATATATGGAACAGTATCCAGCATAT	2528
QY	904	GGTTTCATCTCCTACCTAGAGGCGAGCGGACATATTTCTACCGGACTACGAGAGTGG	963
Db	2527	GGTTTCATATTGACATATGAGGCGAGCGGACGTAATATTTCTACAGGACTTTGAGGAATGG	2468
QY	964	CTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGAGGAGGAC	1023
Db	2467	CTGAACAAGGATAAGCTCAATTAACCTCATTTGGATCCATGATCATTTGGCAGGAGGAGC	2408
QY	1024	ACCGACATAGTCTACTACGATAACGATCACTCTTCGTGAGGAACGGCTACGCGGAC	1083
Db	2407	ACAACAATGTCTACTACGACACCGATGAGCTCATATTTGTGAGAAATGGAGATTCAGA	2348
QY	1084	RAGCCGGGCTTATAACCTACATCAACCTAGCTCGACAGGCGCGGAGGTGGTTTAT	1143
Db	2347	AGGCTTGGGCTTATAACCTTACATTAACCTTAGCCCTAATCTGGTGGTGGGTATAC	2288
QY	1144	GTGCCGAAGTTTCGCGGGCGCTGTCAACGAGTATACCTGTAACCTCGAGGCTGGGTA	1203
Db	2287	GTTCCAAAGTTTCGAGGGGCTTGTTATTATGATGATACCTGGAACCTTAGGAGATGGTA	2228
QY	1204	GACAAGTACGCTACTCAAGCGGCTGGTCTATCTGAACTCAGCTCAGCTTACGACCTGCC	1263
Db	2227	GATAAAGAGTAGTAGTAGCGGATGGGTATACCTAGAGGACCACTCAGCATCCAGCT	2168
QY	1264	AACGGGCAAGTATGGCTACTCCGTGTGGAGTACTCGCGGGTGGGCTGA	1311
Db	2167	RACGCTACTATGGGTACTCCGTATGGAGTTATTTGGTGTGGGTGA	2120

RESULT 15  
AE010170  
LOCUS

DEFINITION	Pyrococcus furiosus DSM 3638, section 45 of 173 of the complete genome.
ACCESSION	AE010170 AE009950
VERSION	AE010170.1 GI:18892445
KEYWORDS	
SOURCE	Pyrococcus furiosus DSM 3638
ORGANISM	Pyrococcus furiosus DSM 3638 Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; Pyrococcus.
REFERENCE	1 (bases 1 to 10643) Maeder, D.L., Weiss, R.B., Dunn, D.M., Cherry, J.L., Gonzalez, J.M., DiRuggiero, J. and Robb, F.T. Divergence of the hyperthermophilic archaea Pyrococcus furiosus and P. horikoshii inferred from complete genomic sequences Genetics 152 (4), 1299-1305 (1999)
JOURNAL	10430560
MEDLINE	99359404
PUBMED	10430560
REFERENCE	2 (bases 1 to 10643) Robb, F.T., Maeder, D.L., Brown, J.R., DiRuggiero, J., Stump, M.D., Yeh, E.K., Weiss, R.B. and Dunn, D.M. Genomic sequence of hyperthermophile, Pyrococcus furiosus: implications for physiology and enzymology Meth. Enzymol. 330, 134-157 (2001)
JOURNAL	21079003
MEDLINE	21079003
PUBMED	11210495
REFERENCE	3 (bases 1 to 10643) Weiss, R.B., Dunn, D.M., Robb, F.T. and Brown, J.R. The complete sequence of the Pyrococcus furiosus genome Unpublished
JOURNAL	4 (bases 1 to 10643) Weiss, R.B. Direct Submission Submitted (12-FEB-2002) Human Genetics, University of Utah, 20 South 2030 East, Salt Lake City, UT 84112, USA
FEATURES	Location/Qualifiers 1..10643 /organism="Pyrococcus furiosus DSM 3638" /mol_type="genomic DNA" /strain="DSM 3638" /db_xref="taxon:186497" complement (68..1261) /genes="PF0471" complement (68..1261) /genes="PF0471" /note="Function Code: 16.1 Conserved Hypothetical" /codon_start=1 /translation="11 /product="hypothetical protein" /protein_id="AAL80595.1" /db_xref="GI:18892446" /translation="MIIPILGILVILTSYTDIKTSYIYEBRFPFSLSKWCKR KGVNKGPIPIVEAGTLYVFFGLKNEIDLAFSGIIGLLGFAVGLLVYVGGW ASGDVLLAFAFSLIPYAPESAKYKPAVAVLPLNATILFNSILLIPELLVYAFAG LIAKKLSLLKELFYKGINRVFVSLWIMFSTLFWVISINIGLSLHLLRWLFFLL IAFLKIRVIGDILGVLVYVLYLVGFSSLYFFLRLLVILLYAFKLLSSVKMRREV LTKRVREIKGEVIGERILEVGGKIVDRKDFEKLNEVLKSGKFEQISGEIAGF SVEGLTKEQIEELKLVSEGLENEFLVRKAMPFALFLGLVLTSYFFGDLWILK VSGMV" complement (1301..1438) /genes="PF0472" complement (1301..1438) /genes="PF0472" /note="Function Code: 16.4 Hypothetical" /codon_start=1 /translation="11 /product="hypothetical protein" /protein_id="AAL80596.1" /db_xref="GI:18892447" /translation="MKRAQTAEYLLMLAAVLILVAIVLVVVVDSMTLTNYVBSQK S" 1492..2172 /genes="PF0473" 1492..2172

```

/genes="PF0473"
/notes="Function Code: 13.2 Translation: Degradation of
proteins, peptides, and glycopeptides; (gcp-2)"
/codon_start=1
/transl_table=11
/product="o-sialoglycoprotein endopeptidase"
/protein_id="AAL80597.1"
/db_xref="GI:18892448"
/transl_table=11
/translation="MRGELMKLIKQGAELKYLADFSLEYDLPDIKIVVERSKRY
RIPEDIKERETREARILHAKAGVNVYFVETKMIIMWIEFIEGERKELL
EKLPIDIERIEIKETGRVIGKHEAGIVHGDITNSMIMRDKGIYILDFGLAEFDDTL
EQAQVDLHLRAMESTHYSWVEGFKAVLRGYEVRGKREIEEKIKIELGRY
RERSMTS"
complement (2162..3850)
/genes="PF0474"
complement (2162..3850)
/genes="PF0474"
/notes="Function Code: 11.1 Replication: DNA replication,
restriction, modification, recombination, and repair;
(mutS)"
/codon_start=1
/transl_table=11
/product="DNA mismatch repair protein"
/protein_id="AAL80598.1"
/db_xref="GI:18892449"
/transl_table=11
/translation="MKLRGDAREVYKRLSLRSMIKLGEARTFLKKEPTSDREIEI
KQNYLKEGLKQVRDDELYLLSIRPFRREFPHDRILLVSDSEVAEKLDLCPVT
SPSEEDYPLLLSTIGYIEVEVKPSHTAPELYIIPWENEDVLEALSVPFGGAAD
KLISUKIBETIFKMEILENDELIIIEKEKELNKKIEKLERFKLTISGRDLVEFMK
ALRAGLELYLFKFSALNDEIIEINKAEKESIDVIGSVEIFPDPFVPPVPEQIEA
LKELEERFKIEFYLSRSTVEIKLPHLOKKEEIOKAYELYFLVWVKFTRDFVPE
IVEEGIFTEGNLFTENPQVSVFYGSKYFNPFGVEANIVILTGANSGKISLEL
ISOVILAHGMPVPAKAAFTVLDBIIFPKKRSVYGAGFETSLKGLNVAIKGKGK
KLILDEFESIPEGAAXITLAELLKIAVEKGFVIVSHLGEDLKRIPIPARVDIE
AKGLENLNIIVDRQPKFVIGRSIPELLIVLARKGRGEKQWNNRILKKFRK"
/genes="PF0475"
complement (3847..4677)
/genes="PF0475"
complement (3847..4677)
/notes="Function Code: 13.6 Translation: Translation
factors; (eIF2B)"
/codon_start=1
/transl_table=11
/product="translation initiation factor eIF-2b, subunit
delta"
/protein_id="AAL80599.1"
/db_xref="GI:18892450"
/transl_table=11
/translation="MLPEKVVKILPEMKNEKTIGASWMAKGAAPILLSBELDSTL
EGEIIELKREILEINSMASLYNLAMFIPITNDREVVKLRAEFIKRAEEAKETASI
GAQIDSGDVIITHSSAVFELFKTAGRGKQFVKVILTESAPDEGLYLAKALQDES
IEVEIITDQGLFKADALVAGDVTGDKGVVKNKAGTYLLAISCVSESVFFVYAA
ETVFKHQKITSEVLERPLVREGSRVENVLFDITPWKPFIRGIIITELGIIILPPDMI"
complement (4677..5435)
/genes="PF0476"
complement (4677..5435)
/genes="PF0476"
complement (4677..5435)
/notes="Function Code: 16 Unknown:"
/codon_start=1
/transl_table=11
/product="competence-damage inducible protein, putative"
/protein_id="AAL80600.1"
/db_xref="GI:18892451"
/transl_table=11
/translation="MPAEIITVGDPELLTGNTVNSNAHAKKLTERRGLVVKRITVGD
DVSDIRSVILESUNRPEDVLIISGGIGPHTDDVTWVAVALAKGELVCEKLEIKE
FYRLRYEGLIDDPDTAKNLPKAGVILPKNSVGAAGAYLEYGRTKIFVLPGM
PREMKAMLEEVLPRIKSKFKVQKYLAEITDESKLAPLEETIKIFDVKHSPPKGF
KFGIIGIIPAOEDDTIKKAVEFMKSGIEFREGW"
5547..6968
/genes="PF0477"
complement (5547..6968)
/genes="PF0477"
complement (5547..6968)
/notes="Function Code: 5.1 Central Intermediary Metabolism:

```

```

Degradation of polysaccharides"
/codon_start=1
/transl_table=11
/product="alpha-amylase"
/protein_id="AAL80601.1"
/db_xref="GI:18892452"
/transl_table=11
/translation="MIHYETLVYGGDHVNIKKLTPLLTLLFFILASPVSAKYLEL
EEGVMQAFYWDVPGGIGWHDHRSKIPEWYEGASIAWLPPSGMGGYSGMYDVP
YDFTDGEVYQKGTVEFRSGSEELVRLIQTAHAGIKVIADVINHRAGGLEWNPPE
VGTITDGEVSKVAGSKITANYLDHFNELHCCDEGTFGFPDICHKKEWQYKLNKS
ESVAAVLSIGFDGFRDYPYKYGAVWVRDLMWGMWAGVAFKATFVANDHLLIWNK
GAKVDFPFLYKDEAFDNNNIPALVYALQNTVVRDPPFKAVTFVANDHLLIWNK
YPAYAFILYEGQVFIYRDFEELNKKDLINLIWHDHLAGGSTIVYDNDLFIY
RNGDSRRPGLITYINLSPNNVGRWVYVPRFAGACIHEYTGNLGGWVKRVDSSGYYL
EAPHPDPANGYCYGYSVYSGVCG"
complement (6894..8966)
/genes="PF0478"
complement (6894..8966)
/genes="PF0478"
complement (6894..8966)
/notes="Function Code: 5.1 Central Intermediary Metabolism:
Degradation of polysaccharides"
/codon_start=1
/transl_table=11
/product="alpha-amylase"
/protein_id="AAL80602.1"
/db_xref="GI:18892453"
/transl_table=11
/translation="MVDRFYDRNPTNNEPFDPEKKNRYLYMGDDIEGIERLDYIES
LGVMWLSPLNDNINRMAGSAPYHYWPRDFKRIDEHFGTWEDFRRLVEEAKKRG
CIIVDVVPHNSPATDEFGALYDNGTLVNTYVEDKKNATNPYTASLENIYHNGNI
NDMGQLYKXANFLGLADFNQNDVDNLYKEGALFVNGACGFRIDAKHIELGWL
ETIYLDIQSDPEFIYGEIFANTPKDFTDLYEFYRYSNVSSLNIPRESIARFPA
YGSFQOLAKMLEEYLSVFPYKQNLFDSDHDLVFLNNPNKORYHNAAGLWMLP
GIVYIYDSELYSKEGSDPYNNPMVDNSTKAAEIRKLSLRKVDALAYSDP
RTVVDYNTWIFERKFGSHKILVANKGPKNITISLNTDSTYDIDIEGAILKVEG
QGIKLPYGFYFVHEBEOKTPLIGITPYIAQPGQKILLAGAGLNINIKYIGGR
ARIIEKEENILVEPEIKTMNAWIPIWVNGTRSNVKRLRYSYNSNDIPALIVLQGN
YTGLVWKNIPELSEPRLLASPTGHYFAVPLPNKFTVQLYKGLWEPLOPTNV
TLIGIGNKIVILNEAAPETPVCGFVAVFALLPLLRKKKKSPNTTITPYGVPIVAV
SWIVRWCL"
complement (9137..10534)
/genes="PF0479"
complement (9137..10534)
/genes="PF0479"
complement (9137..10534)
/notes="Function Code: 16.4 Hypothetical"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL80603.1"
/db_xref="GI:18892454"
/transl_table=11
/translation="MPKKKLLVAVLFLILLSLGRRYLGAFFKPKLOFILVDENGT
PLKISKNVQVQIDALVPTDEVYKTYIGTILKPKTILKFWDSGNLIKTSVADQKIQ
SVISKWVEQSEIEGFSVLSVWVLDYENKALRGRTIYNYEIKDMRKFKKIYKVN
HLPPSELPSNRNTASIQSGSNKQFYKWKVDWSRWRNDYIKIPILIVDNKYSGLIN

```

Query Match	Score	DB 1	Length	10643
Best Local Similarity	75.4%	Pred. No. 1.4e-168		
Matches	986	Conservative	0	Mismatches 322; Indels 0; Gaps 0
QY	4	GCCAAAGTACTCCGAGCTGGAAAAGGCGGGGTATATGATGAGGGGTCTCTACTGGAGCTG	63	
Db	5661	GCAAAATACTTGGAGCTTGAAGAGGGAGGATTTATATGCAAGCATTTCTATTGGGATGT	5720	
QY	64	CCTTCAGGAGGAATATGCTGGGACACAAATACGGCAGAGATACCGAGTGGTACGATGCC	123	
Db	5721	CCAGGGGAGGAATTTGGTGGATCATATAGATCGAAGATTCTGAATGATGAGCT	5780	
QY	124	GGAATCTCCGCAATATCGATTCCCCCGGCGAGCAAGGGCATTGGCGCGCCCTATTCGATG	183	
Db	5781	GGAATCTCTGCAATATGGTACCTCCACCAAGCAAGGGGATGATGGAGGATATTCATG	5840	
QY	184	GGCTACGACCCCTACGACTTCTTTCCTCGGTGAGTACCGACAGAGGAAACGCTAGAG	243	
Db	5841	GGCTACGATCCCTATGATTACTTTGATCTCGGAGTACTTACCAGAGGGAATCTGATAG	5900	

Search completed: June 29, 2004, 10:43:39  
Job time : 5333 secs

Qy 244 ACGCGCTTGGCTCCAAAGCAGGAGCTCGTGAAACATGATAAACCAGCGCCACGCGCTATGGC 303  
Db |||||  
5901 ACGCGCTTGGATCAAAAGAGAACTAGTGAGATTGATACAAACTGCCCATGCCCTATGGA 5960  
Qy 304 ATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGGAAAC 363  
Db |||||  
5961 ATAAGGTAATCGCCGATGATGTTATTAACACACAGGCTGGTGTGACCTAGATGGAAC 6020  
Qy 364 CCCTTCGTGAACGACTATACCTGACCGACTTCTCAAAGGTCGGTCCGGTAAATACAG 423  
Db |||||  
6021 CCCTTCGTGGAGATTACATGACACAGACTTTCTAAAGTTGCCTCAGGGAATATACA 6080  
Qy 424 GCCAATACCTCGACTTCCACCCGAAAGAGCTCCATCGCGGCGATTCCGGAAACATTTGGA 483  
Db |||||  
6081 GCTAACTATCGGACTTCCATCCAAAGAGCTTCATTGTTGTGACGAAGAAACCTTTGGA 6140  
Qy 484 GGCTATCCCGACATATGCCAGCAAGAGCTGGACCCAGTACTGGCTCTCGGCCAGCCAG 543  
Db |||||  
6141 GGATTCAGATATATGTCATCAACAGAGTGGATCAGTACTGGCTATGGAAGACAAAT 6200  
Qy 544 GAGAGCTACGGGGATATCTCAGAGAGCATCGGCATCGATGCTGGCGCTTCGACTACGTC 603  
Db |||||  
6201 GAGAGTTATGCTGCTTATTTAAGAGCATAGGATTTGATGGTTGGAGATTGACTATGTT 6260  
Qy 604 AAGGCTATGCTCCCTGGTCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCGGTT 663  
Db |||||  
6261 AAGGCTATGAGCTTGGGTTGTGAGAGCTGGCTTAATTTGGTGGGAGGTTGGGCAATT 6320  
Qy 664 GGAGAGTACTGGGACACAAAGCTCGAAGCTGTTCTCAACTGGGCAATCTCGAGCGGTGCC 723  
Db |||||  
6321 GGAGAGTACTGGGACACAAATGTAGTGCACACTAAGCTGGGCATATGAGAGTGTGCA 6380  
Qy 724 AAGGCTTTGACTTGGCCCTCTACTACAGATGGATGAGGCTTTGACACAAACAACTT 783  
Db |||||  
6381 AAGGCTTTGACTTGGCCCTCTACTATAAATGGATGAGCAATTTGACAACTT 6440  
Qy 784 CCAGCGCTCGTCTCTGCCCTTCAGAAAGCGGCGAGACTGTTGTCTCCCGCGACCCGTTCAAG 843  
Db |||||  
6441 CCAGCATTTAGTCTATGCCCCCTACAAACGACAACTGTAGTTTCGAGAGATCCATTTAAG 6500  
Qy 844 GCGGTAACCTTTGTAGCAAAACACAGACCGATATATCTGGAACAGTATCCAGGCTAC 903  
Db |||||  
6501 GCAGTAACCTTTGTTGCCAATCATGACACAGATATAATATGGAACAAAGTATCCAGCATAT 6560  
Qy 904 GGGTTTCATCTCAGCTAGAGGCGGCGAGCAATATTTCTACCGGACTACGAGGAGTGG 963  
Db |||||  
6561 GGGTTTCATATTGACATATGAGGACAGCCAGTATATTTCTACAGGACTTTGAGGAATGG 6620  
Qy 964 CTCACAAAGGATPAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGGAGGAAGC 1023  
Db |||||  
6621 CTGAACAAGGATAGCTAATTAACCTCATTTGGATCCATGATCATTTGGCAGGAGGAAGC 6680  
Qy 1024 ACCGACATAGTCTACTAGTAACGATGAACCTCATCTTCTCGAGAACGGCTACGGGAC 1083  
Db |||||  
6681 ACAACAAATGCTCTACTAGCAACCATGAGTCAATTTTGAGMAATGGAGATCTAGA 6740  
Qy 1084 AAGCGGGGCTTTAACTTACCTTACATCAACCTAGGCTCGAGCAAGCGCGGAAGGTGGTTTAT 1143  
Db |||||  
6741 AGGCTTGGGCTTTAACTTACATTTACATTTAGSCCTTAACCTGGTTGGTGGGTATAC 6800  
Qy 1144 GTGCGGAAGTTCCGGGGCGGTGATCCACGAGTATAGTGAACCTCGGAGGCTGGGTA 1203  
Db |||||  
6801 GTTCCAAAGTTTGCAGGGGCTTGTATTATGAATACACTGGAACCTTAGGAGGATGGGTA 6860  
Qy 1204 GACAGTAGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
Db |||||  
6861 GATAAAGAGTAGATAGTAGCGGATGGGTATACCTAGAGGCACCACTCAGATCCAGCT 6920  
Qy 1264 AACGGGCACTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGCTGA 1311  
Db |||||  
6921 AACGGCTACTATGGGTACTCCGTATGGAGTTATTGTGGTGTGGGTGA 6968



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 09:04:55 ; Search time 573 Seconds  
(without alignments)  
9719.699 Million cell updates/sec

Title: US-10-081-739A-1  
Perfect score: 1311  
Sequence: 1 atggccaagtactccgagct.....gctactgcgggtggcgctga 1311

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_29Jan04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1980s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002s:\*
  - 7: Geneseqn2003as:\*
  - 8: Geneseqn2003bs:\*
  - 9: Geneseqn2003cs:\*
  - 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1311	7 ABS55928	DNA encod
2	1311	100.0	1311	7 ABS223371	Nucleotid
3	1239	94.5	1311	7 ABX08413	DNA encod
4	1231	93.9	1311	7 ABX08445	DNA encod
5	1192.6	91.0	1311	7 ABX08428	DNA encod
6	1186.2	90.5	1311	7 ABX08429	DNA encod
7	1183	90.2	1311	7 ABX08424	DNA encod
8	1171.8	89.4	1311	7 ABX08416	DNA encod
9	1160.6	88.5	1311	7 ABX08440	DNA encod
10	1155.8	88.2	1311	7 ABX08437	DNA encod
11	1147.8	87.6	1311	7 ABX08432	DNA encod
12	1147.8	87.6	1311	7 ABX08432	DNA encod
13	1146.2	87.4	1311	7 ABX08412	DNA encod
14	1146.2	87.4	1311	7 ABX08439	DNA encod
15	1144.6	87.3	1311	7 ABX08435	DNA encod
16	1143	87.2	1311	7 ABX08441	DNA encod
17	1141.4	87.1	1311	7 ABX08419	DNA encod
18	1139.8	86.9	1311	7 ABX08417	DNA encod
19	1138.2	86.8	1311	7 ABX08415	DNA encod
20	1130.2	86.2	1311	7 ABX08430	DNA encod
21	1120.6	85.5	1311	7 ABX08411	DNA encod
22	1119	85.4	1311	7 ABX08427	DNA encod
23	1117.4	85.2	1311	7 ABX08421	DNA encod

24	1117.4	85.2	1311	7 ABX08425	DNA encod
25	1114.2	85.0	1311	7 ABX08420	DNA encod
26	1112.6	84.9	1311	7 ABX08414	DNA encod
27	1111	84.7	1311	7 ABX08442	DNA encod
28	1102.6	84.1	1299	7 ABX08448	DNA encod
29	1096.6	83.6	1311	7 ABX08426	DNA encod
30	1095	83.5	1311	7 ABX08423	DNA encod
31	1089.8	83.1	1299	7 ABS55932	DNA encod
32	1089.8	83.1	1299	7 ABX08447	DNA encod
33	1085	82.8	1386	7 ABS55933	DNA encod
34	1085	82.8	1386	7 ABX08449	DNA encod
35	1078.6	82.3	1311	7 ABX08433	DNA encod
36	1077.4	82.2	1311	7 ABX08431	DNA encod
37	1074.2	81.9	1311	7 ABX08436	DNA encod
38	1042.2	79.5	1378	7 ABX08513	DNA encod
39	1040.8	79.4	1386	7 ABS55931	DNA encod
40	1040.8	79.4	1386	7 ABX08450	DNA encod
41	1034.4	78.9	1575	2 AAT94206	AA94206 AEP11a c
42	1034.2	78.9	1311	7 ABX08418	DNA encod
43	1031	78.6	1311	7 ABX08443	DNA encod
44	1003.4	76.5	2705	3 AA236889	AA236889 Genomic D
45	1000.2	76.3	1299	7 ABX08446	DNA encod

ALIGNMENTS

RESULT 1  
ABS55928 standard; DNA; 1311 BP.  
XX  
AC ABS55928;  
XX  
DT 20-JAN-2003 (first entry)  
XX  
DE DNA encoding alpha amylase polypeptide #1.

Alpha amylase; gene; ds; starch liquefaction; starch linkage hydrolysis;  
starch breakdown catalysis; high-maltose syrup; high-glucose syrup; rice;  
germinated rice; corn; barley; wheat; legume; sweet potato; anti-staling;  
textile desizing; lignocellulosic fibre treatment; enzymatic de-inking;  
recycled paper pulp; alpha-glucosidic linkage; dough; crumb firming;  
crumb elasticity; palatability; improved flavour.

Unidentified.

Key Location/Qualifiers  
CDS 1..1311  
FT /\*tag= a  
FT /product= "Alpha amylase #1"

MO200268597-A2.

06-SEP-2002.

21-FEB-2002; 2002WO-US005538.

21-FEB-2001; 2001US-0270495P.

21-FEB-2001; 2001US-0270496P.

14-MAY-2001; 2001US-0291122P.

(DIVE-) DIVERSA CORP.

Callen W, Richardson T, Frey G, Miller C, Kazaoka M, Mathur EJ;

Short JW;

WPI; 2003-018657/01.

P-PSDB; ABG70531.

Novel alpha amylase polypeptide, useful e.g. for starch liquefaction,

textile desizing, treating lignocellulosic fibers and for preparing dough

or baked product, has increased activity and stability.

XX

PS Claim 1; Fig 7; 147pp; English.

XX The invention relates to a purified alpha amylase polypeptide and the  
 CC polynucleotide encoding it. The polypeptide is useful for liquefying a  
 CC starch containing composition, hydrolysing a starch linkage, catalysing  
 CC the breakdown of a starch and producing high-maltose or high-glucose  
 CC syrup or mixed syrup, where the starch is from rice, germinated rice,  
 CC corn, barley, wheat, legumes or sweet potato. The polypeptide is also  
 CC useful for washing an object, textile desizing, treating lignocellulosic  
 CC fibres for improving fibre properties, enzymatic de-inking of recycled  
 CC paper pulp and modifying small molecules by biocatalytic reaction. The  
 CC polypeptide is further useful for increasing the flow of production  
 CC fluids from a subterranean formation by removing a viscous, starch-  
 CC containing, damaging fluid formed during production operations and found  
 CC within the subterranean formation which surrounds a completed well bore,  
 CC by pumping an enzyme treatment comprising the polypeptide and an aqueous  
 CC fluid to a desired location within the well-bore which is effective to  
 CC attack the alpha-glucosidic linkages in the starch-containing fluid. The  
 CC alpha amylases have increased activity and stability at increased pH and  
 CC temperature and can be used in preparing dough to provide an improved  
 CC anti-staling effect as measured by less crumb firming, retained crumb  
 CC elasticity, improved slice-ability, improved palatability or improved  
 CC flavour. This sequence represents DNA encoding an alpha amylase of the  
 CC invention

SQ Sequence 1311 BP; 318 A; 364 C; 376 G; 253 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 7; Length 1311;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCAACTACCGAGTGGAAAGGGCGGGTCAATACGAGCGCTTCTACTGGGAC 60  
 DB 1 ATGGCCAACTACCGAGTGGAAAGGGCGGGTCAATACGAGCGCTTCTACTGGGAC 60

QY 61 GTCCCTTCAGGAGGAATATGGTGGGACACAAATACGGCAGAGATACCGGAGTGATACGAT 120  
 DB 61 GTCCCTTCAGGAGGAATATGGTGGGACACAAATACGGCAGAGATACCGGAGTGATACGAT 120

QY 121 GCCGGAATCTCCGCAATATGATATCCCGGGGAGCAGAGGATACGGCGGCGCTATTTCG 180  
 DB 121 GCCGGAATCTCCGCAATATGATATCCCGGGGAGCAGAGGATACGGCGGCGCTATTTCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACACAGAGGCAACGGTA 240  
 DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACACAGAGGCAACGGTA 240

QY 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATATAACACCGCCCAACGCTAT 300  
 DB 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATATAACACCGCCCAACGCTAT 300

QY 301 GGCATGAAGTAAATAGCCGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGG 360  
 DB 301 GGCATGAAGTAAATAGCCGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTGCGCGGTAAATAC 420  
 DB 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTGCGCGGTAAATAC 420

QY 421 ACGGCCAACTACCTGACTTCCACCGAAGAGCTCCATCGGGGATTCGGGACATTT 480  
 DB 421 ACGGCCAACTACCTGACTTCCACCGAAGAGCTCCATCGGGGATTCGGGACATTT 480

QY 481 GGAGGCTATCCCGACATATGCGACCAAGAGCTGGGACCACTACTGGCTCTGGGCCAGC 540  
 DB 481 GGAGGCTATCCCGACATATGCGACCAAGAGCTGGGACCACTACTGGCTCTGGGCCAGC 540

QY 541 CAGGAGGCTACCGGGCATATCTCAGAGGATCGGATCGATGCTTGGCGCTTCGACTAC 600  
 DB 541 CAGGAGGCTACCGGGCATATCTCAGAGGATCGGATCGATGCTTGGCGCTTCGACTAC 600

QY 601 GTCAAGGGCTATCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
 DB 601 GTCAAGGGCTATCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660

Db 601 GTCAAGGGCTATCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACAGCTCGACCGTGTCTCAACTGGGCATATCTGAGCGGT 720  
 Db 661 GTTGGAGAGTACTGGGACACCAACAGCTCGACCGTGTCTCAACTGGGCATATCTGAGCGGT 720  
 QY 721 GCCAAGGCTTTTGAATCTGGCCCTCTACTACAAGATGATGAGCCCTTTGACAAACAAAAC 780  
 Db 721 GCCAAGGCTTTTGAATCTGGCCCTCTACTACAAGATGATGAGCCCTTTGACAAACAAAAC 780  
 QY 781 ATTCCAGCGCTGCTCTGCGCTTCAAGAGCGGACAGCTGTGTCTCCCGCCGACCCGTTTC 840  
 Db 781 ATTCCAGCGCTGCTCTGCGCTTCAAGAGCGGACAGCTGTGTCTCCCGCCGACCCGTTTC 840  
 QY 841 AAGGCGGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAAACAAAGTATCCAGCC 900  
 Db 841 AAGGCGGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAAACAAAGTATCCAGCC 900  
 QY 901 TAGCGGTTCATCTCACTACGAGGGCCGACGACATATTTACCGGACTACGAGGAG 960  
 Db 901 TAGCGGTTCATCTCACTACGAGGGCCGACGACATATTTACCGGACTACGAGGAG 960  
 QY 961 TGGCTCAACAAAGGATAAGCTCAAGAAACCTCATCTGGATACATGAGAACTCTCGCGGAGGA 1020  
 Db 961 TGGCTCAACAAAGGATAAGCTCAAGAAACCTCATCTGGATACATGAGAACTCTCGCGGAGGA 1020  
 QY 1021 AGACCGGATAGTCTACTACGATACGATGAATCATCTTCTGTCAGGACCGCTACGGG 1080  
 Db 1021 AGACCGGATAGTCTACTACGATACGATGAATCATCTTCTGTCAGGACCGCTACGGG 1080  
 QY 1081 GACAAGCGGGGCTTATACTACATCAACTAGGCTCGAGCAAGCCGGAAGCTGGGT 1140  
 Db 1081 GACAAGCGGGGCTTATACTACATCAACTAGGCTCGAGCAAGCCGGAAGCTGGGT 1140  
 QY 1141 TATGTGCGAAGTTTCGGCGGCGTGCATCCACAGTATACCTGTAACCTCGGAGGCTGG 1200  
 Db 1141 TATGTGCGAAGTTTCGGCGGCGTGCATCCACAGTATACCTGTAACCTCGGAGGCTGG 1200  
 QY 1201 GTAGACAAGTACGCTACTCAAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 Db 1201 GTAGACAAGTACGCTACTCAAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 QY 1261 GCCAACGGGCGAGTATGCTACTCCGTGTGAGTACTTGGGGGTGGGCTGA 1311  
 Db 1261 GCCAACGGGCGAGTATGCTACTCCGTGTGAGTACTTGGGGGTGGGCTGA 1311

RESULT 2  
 ABZ23371  
 ID ABZ23371 standard; DNA; 1311 BP.  
 XX  
 AC ABZ23371;  
 XX  
 DT 07-APR-2003 (first entry)  
 XX  
 DE Nucleotide sequence of an alpha-amylase.  
 XX  
 KW Alpha-amylase; enzyme; detergent; baking; beverage; fuel; ethanol; gene;  
 XX ss.  
 XX  
 OS Bacillus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1311  
 FT /\*tag= a  
 FT /product= "alpha-amylase"  
 XX  
 XX WO200292802-A1.  
 XX  
 XX 21-NOV-2002.  
 XX  
 XX 14-MAY-2002; 2002WO-US015422.  
 XX





XX 21-FEB-2002; 2002WO-US005068.  
 XX PF  
 XX XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX XX  
 PA (DIVE-) DIVERSA CORP.  
 XX XX  
 XX Callen W, Richardson T, Frey G;  
 XX WPI; 2003-018656/01.  
 DR P-FSDS; ABU03042.  
 DR XX  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 PT XX  
 XX Claim 1; Fig 16; 30pp; English.  
 PS XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing,  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 CC XX  
 SQ Sequence 1311 BP; 312 A; 367 C; 375 G; 257 T; 0 U; 0 Other;

Query Match 94.5%; Score 1239; DB 7; Length 1311;  
 Best Local Similarity 96.6%; Pred. No. 0;  
 Matches 1266; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATGGCCAACTCTCCGAGCTGGAAAGGGCGGGGTCAATATGACGGCGTCTACTGGGAC 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1 ATGGCCAACTCTCCGAGCTGGAAAGGGCGGGGTCAATATGACGGCGTCTACTGGGAC 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 GTGCCCTTCAGGAGGAATATGTTGGGACACAATACGGGCGAAGATACCGAGTGTACGAT 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 GTCCCAAGTGGAGGAATCTGGTGGGACACCATCAGGAGCAGATACCGAGTGTACGAG 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 GCGGGAATCTCGCAATATGATATGATTCGCCCGGAGCAGGAGGATGGCGGCGCTATTTCG 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 GCGGGAATATCGGCCATTTTGGATTTCCCGGCAAGCAAGGGCATGGGCGGCGCTATTTCG 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 ATGGGGTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 ATGGGGTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 241 GAGACCGGCTTTGGCTTCCAGCAGGAGCTCGTGAACATGATATAAACACCGCCACGCTTAT 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 241 GAGACCGGCTTTGGCTTCCAGCAGGAGCTCGTGAACATGATATAAACACCGCCACGCTTAT 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACCCCGCGCGGTGACCTGGAGTGG 360  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACCCCGCGCGGTGACCTGGAGTGG 360  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTCGCGGTAAATATAC 420  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTCGCGGTAAATATAC 420  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 421 ACGGCCAACTACCTCGACTTCCACCAGAACGAGCTCCATGCGGGCGATTCCGGAAACATTT 480  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 421 ACGGCCAACTACCTCGACTTCCACCAGAACGAGCTCCATGCGGGCGATTCCGGAAACATTT 480  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 481 GGAGGCTATCCCGACATATGCCACGACAAGAGCTGGGACAGTACTGGCTCTGGGGCAGC 540  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 481 GGAGGCTATCCCGACATATGCCACGACAAGAGCTGGGACAGTACTGGCTCTGGGGCAGC 540  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGCATCGATCGCTTGGCGCTTCGACTAC 600  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGCATCGATCGCTTGGCGCTTCGACTAC 600  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 601 GTCAAGGGCTATGCTCCCTGGTCTCAAGGCTGGCTGAATGGTGGGAGGCTGGGG 660  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 601 GTCAAGGGCTATGCTCCCTGGTCTCAAGGCTGGCTGAATGGTGGGAGGCTGGGG 660  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 661 GTTGGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATATCGAGCGGT 720  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 661 GTTGGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATATCGAGCGGT 720  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 721 GCCAAGGCTTTGACTTCGCTCTACTACAGATGGATGAGGCTTTGACAAACAAAC 780  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 721 GCCAAGGCTTTGACTTCGCTCTACTACAGATGGATGAGGCTTTGACAAACAAAC 780  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 781 ATTCCAGCGCTCGTCTCTGCGCTTCAAGAGGCGACAGTGTCTCTCCCGGACCCGCTTC 840  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 781 ATTCCAGCGCTCGTCTCTGCGCTTCAAGAGGCGCGACAGTGTCTCTCCCGGACCCGCTTC 840  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 841 AAGGCGGTAACTTTGAGCAACACGACGCGATATAATCTGGAAACAGTACCTTGTCT 900  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 841 AAGGCGGTAACTTTGAGCAACACGACGCGATATAATCTGGAAACAGTACCTTGTCT 900  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 901 TAGCGCTTCATCTCCTACCTACGAGGCGCGACGACATATTTACCGGACTACGAGGAG 960  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 901 TAGCGCTTCATCTCCTACCTACGAGGCGCGACGCGCTCATATTTCTACCGGACGAGGAG 960  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 961 TGGCTCAACAAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGA 1020  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 961 TGGCTCAACAAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGA 1020  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1021 AGACCGGACTACTACTACGATACGATGAGTCACTCTTCTGAGGAGCGGCTACGGG 1080  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1021 AGACCGGACTACTACTACGATACGATGAGTCACTCTTCTGAGGAGCGGCTACGGG 1080  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1081 GACAGCCGGGGTATAACCTACATCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGGTT 1140  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1081 GACAGCCGGGGTATAACCTACATCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGGTT 1140  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1141 TATGTGCCGAAGTTCCGGGGCGGTGCTCCAGAGTATCTGGTAACTCGAGGCTGG 1200  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1141 TATGTGCCGAAGTTCCGGGGCGGTGCTCCAGAGTATCTGGTAACTCGAGGCTGG 1200  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1201 GTAGCAAGTACGCTTACTCAAGCGCTGGGTCTATCTGGAAGCTCCAGCTTACGACCT 1260  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1201 GTAGCAAGTACGCTTACTCAAGCGCTGGGTCTATCTGGAAGCTCCAGCTTACGACCT 1260  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1261 GCCAAGGGGAGTATGGCTTACTCCGTGTGGAGTACTTGGGGGTGGGCTGA 1311  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1261 GCCAAGGGGAGTATGGCTTACTCCGTGTGGAGTACTTGGGGGTGGGCTGA 1311  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 4

ABX08445

ID ABX08445 standard; DNA; 1311 BP.

XX

AC ABX08445;

XX

DT 21-JAN-2003 (first entry)

XX

DE DNA encoding alpha amylase polypeptide #35.

XX

KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 starch breakdown catalysis; textile desizing; lignocellulosic fibre;

XX	enzymatic de-inking; recycled paper; high-maltose syrup; dough;
XX	high glucose syrup; corn-wet milling process; detergent; baking process;
XX	beverage; oil field; fuel ethanol; brewing process; staling;
XX	starch modification.
XX	Unidentified.
OS	
XX	WO200268589-A2.
PN	
XX	06-SEP-2002.
PD	
XX	21-FEB-2002; 2002WO-US005068.
XX	
PR	21-FEB-2001; 2001US-0270495P.
PR	21-FEB-2001; 2001US-0270496P.
PR	14-MAY-2001; 2001US-0291122P.
XX	
XX	(DIVE-) DIVERSA CORP.
XX	
XX	Callen W, Richardson T, Frey G;
PI	
XX	WPI; 2003-018656/01.
DR	P-FSDB; ABU03074.
XX	
XX	Novel purified polypeptide with alpha-amylase activity, useful e.g. for
PT	liquefying starch, for textile desizing, for treating lignocellulosic
PT	fibers, and for producing high-maltose or high-glucose syrup.
PT	
XX	Claim 1; Fig 16; 30pp; English.
PS	
XX	
XX	The invention relates to a purified polypeptide with alpha-amylase
CC	activity and the polynucleotide encoding it. The polypeptide is useful
CC	for hydrolysing starch linkages, for catalysing the breakdown of a
CC	starch, for modifying small molecules, for liquefying starch, for washing
CC	an object, for textile desizing, for treating lignocellulosic fibers, for
CC	improving fibre properties, for enzymatic de-linking of recycled paper
CC	pulp, for producing a high-maltose or high-glucose syrup or a mixed
CC	syrup, and for increasing the flow of production fluids from a
CC	subterranean formation by removing a viscous, starch-containing, damaging
CC	fluid formed during production operations and found within the
CC	subterranean formation which surrounds a completed well bore. The
CC	polypeptide is also useful for preparing a dough or a baked product
CC	prepared from the dough and in corn-wet milling processes, detergents,
CC	baking processes, beverages, oil fields (fuel ethanol), brewing processes
CC	and starch modification in the paper and pulp industry, for removing
CC	starch containing stains from a material and for reducing staling of
CC	bakery products. Sequences ABX08411-ABX08515 represent DNA molecules
CC	encoding alpha amylase polypeptides of the invention
XX	
XX	
SQ	Sequence 1311 BP; 322 A; 362 C; 376 G; 251 T; 0 U; 0 Other;
	Query Match 93.9%; Score 1231; DB 7; Length 1311;
	Best Local Similarity 96.2%; Pred. No. 0;
	Matches 1261; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY	1 ATGGCCAAGTATCTCGAGCTGGAAAGGGCGGGGTCTAATGAGGGTCTTACTGGGAC 60
DB	1 ATGGCCAAGTACCTCGAGCTCGAAGGGCGGGGTCTAATGAGGGTCTTACTGGGAC 60
QY	61 GTGGCTTCAGGAGGAATATGGTGGGACACAAATACGCAGAAAGATACCGAGTGGTACGAT 120
DB	61 GTGGCTTCAGGAGGAATATGGTGGGACACAAATACGCAGAAAGATACCGAGTGGTACGAT 120
QY	121 GCGGGAATCTCGGAATATGGATTCCCGCGAGCAAGGGCATGGGCGCGCGCTATTTCG 180
DB	121 GCGGGAATCTCGGAATATGGATTCCCGCGAGCAAGGGCATGGGCGCGCGCTATTTCG 180
QY	181 ATGGGGCTAGACCCCTCAGACTTTCTTTTGACCTCGTGAGTACACACAGAAAGGGAACGGTA 240
DB	181 ATGGGGCTAGACCCCTCAGACTTTCTTTTGACCTCGTGAGTACACACAGAAAGGGAACGGTA 240
QY	241 GAGACGCGCTTTGGTCCCAAGCAGAGCTCGTGAACATGATAAACAACGCCCAACGCTTAT 300

ID ABX08428 standard; DNA; 1311 BP.

AC ABX08428;

DT 21-JAN-2003 (first entry)

DE DNA encoding alpha amylase polypeptide #18.

XX Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.

XX Unidentified.

XX WO200268589-A2.

XX 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005068.

XX 21-FEB-2001; 2001US-0270495P.

XX 21-FEB-2001; 2001US-0270496P.

XX 14-MAY-2001; 2001US-0291122P.

XX (DIVE-) DIVERSA CORP.

XX Callen W, Richardson T, Frey G;

XX WPI: 2003-018656/01.

XX P-PSDB; AB030357.

XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
PT liquefying starch, for textile desizing, for treating lignocellulosic  
PT fibers, and for producing high-maltose or high-glucose syrup.

XX Claim 1; Fig 16; 301pp; English.

XX The invention relates to a purified polypeptide with alpha-amylase  
CC activity and the polynucleotide encoding it. The polypeptide is useful  
CC for hydrolysing starch linkages, for catalysing the breakdown of a  
CC starch, for modifying small molecules, for liquefying starch, for washing  
CC an object, for textile desizing, for treating lignocellulosic fibers, for  
CC improving fibre properties, for enzymatic de-linking of recycled paper  
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
CC encoding alpha amylase polypeptides of the invention

XX Sequence 1311 BP; 317 A; 373 G; 378 G; 243 T; 0 U; 0 Other;

XX Query Match 91.0%; Score 1192.6; DB 7; Length 1311;

XX Best Local Similarity 94.4%; Pred. No. 0;

XX Matches 1237; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATGGCCCAAGTACTCCAGCTGGAAAGGCGGGGTCATAATGACGCGTTCTACTGGGAC 60

DB 1 ATGGCCCAAGTACTCCAGCTGGAAAGGCGGGGTCATAATGACGCGTTCTACTGGGAC 60

QY 61 GTGCTTCAGGAGGAATATGGTGGGACACAATACGCGAAGATACCGGAGTGGTACGAT 120

DB 61 GTGCTTCAGGAGGAATATGGTGGGACACAATACGCGAAGATACCGGAGTGGTACGAT 120

121 GCCGGAATCTCCGAATATGGATTCCCCCGGCGAGCAAGGGCATGGCGGCGCCTATTTCG 180

121 GCCGGAATCTCCGAATATGGATTCCCCCGGCGAGCAAGGGCATGGCGGCGCCTATTTCG 180

181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAACGGTA 240

181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAACGGTA 240

241 GAGACCGGTTTGGCTTCCAAAGAGGAGTCTGTAACATGATAAACCACCCACGCTAT 300

241 GAGACCGGTTTGGCTTCCAAAGAGGAGTCTGTAACATGATAAACCACCCACGCTAT 300

301 GGCATGAAGGTAAATAGCCGATATAGTCAACACCGCGCGGCGGTGACCTGAGTGG 360

301 GGCATGAAGGTAAATAGCCGATATAGTCAACACCGCGCGGCGGTGACCTGAGTGG 360

361 AACCCCTTCTGTGAACGACTATACCTGGACCGGACTTCTCAAGGTTCGCTCGGTAATATAC 420

361 AACCCCTTCTGTGAACGACTATACCTGGACCGGACTTCTCAAGGTTCGCTCGGTAATATAC 420

421 ACGGCCAATACCTCGACTTCCACCCGAAGAGTCTCATCGGGCGGATTCGGAACATTT 480

421 ACGGCCAATACCTCGACTTCCACCCGAAGAGTCTCATCGGGCGGATTCGGAACATTT 480

481 GGAGGCTATCCCGACATATGCCAGCAAGAGTGGGACAGTACTTGGCTCTGGGCGAC 540

481 GGAGGCTATCCCGACATATGCCAGCAAGAGTGGGACAGTACTTGGCTCTGGGCGAC 540

541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGGCGCTTGTACTAC 600

541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGGCGCTTGTACTAC 600

601 GTCAAGGCTATGCTCCCTGGTCTGCAAGGACTGGCTGAACTGGTGGGAGGCTGGGCG 660

601 GTCAAGGCTATGCTCCCTGGTCTGCAAGGACTGGCTGAACTGGTGGGAGGCTGGGCG 660

661 GTTGGAGGCTACGCGGCATCGGCATCGGCATCGGCATCGGCATCGGCATCGGCATCGGC 720

661 GTTGGAGGCTACGCGGCATCGGCATCGGCATCGGCATCGGCATCGGCATCGGCATCGGC 720

721 GCCAAGGCTTTCGACTTCCGCTCTACTACAGATGGATGAGGCTTTCGACCAACAAAC 780

721 GCCAAGGCTTTCGACTTCCGCTCTACTACAGATGGATGAGGCTTTCGACCAACAAAC 780

781 ATTCCAGCGCTCTCTCTGCTCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 840

781 ATTCCAGCGCTCTCTCTGCTCTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 840

841 AAGCGGTAACCTTTAGCAACACGACACCGATATATCTGGAACAAAGTATCCAGGC 900

841 AAGCGGTAACCTTTAGCAACACGACACCGATATATCTGGAACAAAGTATCCAGGC 900

901 TAGCGGTTCTATCTCTACCTACGAGGCGCCAGCAATATCTTACCGCACTACGAGGAG 960

901 TAGCGGTTCTATCTCTACCTACGAGGCGCCAGCAATATCTTCTCGCACTACGAGGAG 960

961 TGCGTCAACAAAGGATAAGTCTCAAGAACCTCATCTGGATACATGAGAACCTTCCGCGGAGGA 1020

961 TGCGTCAACAAAGGATAAGTCTCAAGAACCTCATCTGGATACATGAGAACCTTCCGCGGAGGA 1020

1021 AGCACCGCATAGTCTTACTACGATACGATGACTCATCTTCTCTAGGAAACCGGTACGGG 1080

1021 AGCACGTGACATCTTTTACTACGACACAGGAGCTGATGATTTCTGTGAGAAACCGGTACGGA 1080

1081 GACAAGCCGGGGTATTAACCTACATCAACCTAGGCTCGAGCAAGCCCGGAAGTGGTT 1140

1081 GACAAGCCGGGGTATTAACCTACATCAACCTAGGCTCGAGCAAGCCCGGAAGTGGTT 1140

1141 TATGTGCGGAGTTCGGGCGGCTGCATCCAGGATATCTGTTAACTCCGAGGCTCG 1200

1141 TATGTGCGGAGTTCGGGCGGCTGCATCCAGGATATCTGTTAACTCCGAGGCTCG 1200

1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

1201 GTGGCAAGTGGTGGACTCAAGCGCTGGGTCTACCTCGAGGCTCTGCGCCAGACCGG 1260  
 1261 GCCAACGGCGAGTAGGCTACTCCGTGTGGAGCTACTGCGGGGTGGGTGA 1311  
 1261 GCCAACGGCCAGTACGGCTACTCCGTGTGGAGCTACTGCGGGGTGGGTGA 1311

RESULT 6  
 ID ABX08429 standard; DNA; 1311 BP.  
 AC ABX08429;  
 XX  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE DNA encoding alpha amylase polypeptide #19.  
 XX  
 KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough; baking process;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200268589-42.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-FEB-2002; 2002WO-US005068.  
 XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 XX  
 PI Callen W, Richardson T, Frey G;  
 XX  
 DR WPI; 2003-018656/01.  
 DR P-FSDB; ABU03058.  
 XX  
 PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX  
 PS Claim 1; Fig 16; 30pp; English.  
 XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 XX  
 SQ Sequence 1311 BP; 312 A; 372 C; 380 G; 247 T; 0 U; 0 Other;

Query Match 90.5%; Score 1186.2; DB 7; Length 1311;  
 Best Local Similarity 94.1%; Pred. No. 0;

Matches 1233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
 QY 1 ATGGCCAAGTACTCCGAGCTGGAAAAAGCGGGTGCATATGCAAGCGTTCTACTGGGAC 60  
 DB 1 ATGGCCAAGTACTCCGAGCTCGAAGAGGGCGGGTGCATATGCAAGCGTTCTACTGGGAC 60  
 QY 61 GTGCCCTTCAGGAGGAATATGTTGGGACACAATACCGCAGAAGNATACCGAGTGGTACGAT 120  
 DB 61 GTGCCCTTCAGGAGGAATATGTTGGGACACAATACCGCAGAAGNATACCGAGTGGTACGAT 120  
 QY 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGGAGCAAGGGCATGGGCGGCGCTATTTCG 180  
 DB 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGGAGCAGGGGTATGAGCGGCGCTATTTCG 180  
 QY 181 ATGGGCTACAGCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGAACCGTA 240  
 DB 181 ATGGGCTACAGCCCTACGATTATTTTGACCTCGGTGAGTACTACCAAGGAACCGGTG 240  
 QY 241 GAGACCGGCTTTGGCTCCCAAGCAGGAGCTCGTGAACATGATAAAACACCCGCCACGCTAT 300  
 DB 241 GAAACGAGGTTGGCTCAAGCAGGAGCTCATAAACATGATAAAACACCCGCCACGCTAT 300  
 QY 301 GGCAATGAAGTAATAGCCGATATAGTATCAACACCGCGCGGGGTGACCTGGAAGTGG 360  
 DB 301 GGCAATGAAGTAATAGCCGATATAGTATCAACACCGCGCGGGGTGACCTGGAAGTGG 360  
 QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTGCGCTCGGTAATAC 420  
 DB 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTGCGCTCGGTAATAC 420  
 QY 421 ACGGCCAACTACCTCGACTTCCACCCGAAACGAGCTCCATGCGGGCGATTCCGGAACATTT 480  
 DB 421 ACGGCCAACTACCTCGACTTCCACCCGAAACGAGCTCCATGCGGGCGATTCCGGAACATTT 480  
 QY 481 GGAGGCTATCCGACATATGCCACGACGAGCTGGACACGAGTCTGGGTCTGGGCGACG 540  
 DB 481 GGAGGCTATCCGACATATGCCACGACGAGCTGGACACGAGTCTGGGTCTGGGCGACG 540  
 QY 541 CAGGAGAGCTAGCGGCATATCTCAGGAGCATCTCGGACATCGATGCGTGGCGTTCGACTAC 600  
 DB 541 CAGGAGAGCTAGCGGCATATCTCAGGAGCATCTCGGACATCGATGCGTGGCGTTCGACTAC 600  
 QY 601 GTCAAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAATGTGTGGGAGGCTGGGCG 660  
 DB 601 GTCAAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAATGTGTGGGAGGCTGGGCG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGACATCTCGAGCGGT 720  
 DB 661 GTTGGAGAGTACTGGGACACCAACGTTGATGCCCTCCCTCGGCTTACTCGAGCGGC 720  
 QY 721 GCCAAGGTCTTTGACTTCGCCCTCTACTACAAGATGGATGAGGCTTTTGACAACAAAAAC 780  
 DB 721 GCCAAGGTCTTCGACTTCCGCTCTACTACAAGATGGATGAGGCTTTTGACAACAAAAAC 780  
 QY 781 ATTCCAGCGCTGCTCTGCCCTTCAGACGCGGACAGCTGTTGTCTCCCGCACCGCTTC 840  
 DB 781 ATTCCAGCGCTGCTCTGCCCTTCAGACGCGGACAGCTGTTGTCTCCCGCACCGCTTC 840  
 QY 841 AAGGCCGTAACTTTTGTAGCAAAACCAACGACCCGATATATCTGGAACAAGTATCCAGCC 900  
 DB 841 AAGGCCGTAACTTTTGTAGCAAAACCAACGATACGATATATCTGGAACAAGTATCCAGCC 900  
 QY 901 TAGCGTTTCATCTCTACCTACGAGGCGCAGCGACAAATATTTCTACCGGACTACGAGGAG 960  
 DB 901 TAGCGTTTCATCTCTACCTACGAGGCGCAGCGACAAATATTTCTACCGGACTACGAGGAG 960  
 QY 961 TGGCTCAACAGGATTAAGCTCAAGAACCTCATCTGATACATGAAACCTCCCGCGAGGA 1020  
 DB 961 TGGCTCAACAGGATTAAGCTCAAGAACCTCATCTGATACATGAAACCTCCCGCGAGGA 1020  
 QY 1021 AGCACCGACATAGTCTACTACGATTAACGATGAACCTCATCTTCGTACGAGAACGCTACGGG 1080  
 DB 1021 AGCACCGACATAGTCTACTACGATTAACGATGAACCTCATCTTCGTACGAGAACGCTACGGG 1080



Db 901 TATGCTTTCATCTCACCTACGAGGCCAGCCGCTCATATCTACCGGACTACGAGGAG 960  
 QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGGAGGA 1020  
 Db 961 TGGCTCAACAGGACAGGTTGAAACCTCATATGGATACACGACCTCGAGGGGGA 1020  
 QY 1021 AGCACCGACATAGTCTACTAGGATAACGATGAACCTCATCTTCGTGAGGACGGCTACGGG 1080  
 Db 1021 AGCACCGACATAGTCTACTAGGATAACGATGAACCTCATCTTCGTGAGGACGGCTACGGG 1080  
 QY 1081 GACAGACCGGGCTTATACCTACATCACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1140  
 Db 1081 GACAGACCGGGCTTATACCTACATCACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1140  
 QY 1141 TATGTCCGAAGTTCGCGGGCGCTGCATCCACGAGTATAGTGAACCTCGAGGCTGG 1200  
 Db 1141 TATGTCCGAAGTTCGCGGGCGCTGCATCCACGAGTATAGTGAACCTCGAGGCTGG 1200  
 QY 1201 GTAGACAAGTACGCTTACTCAAGCGGCTGGTCTATCTCGAAGCTTCAGCTTACGACCT 1260  
 Db 1201 GTAGACAAGTACGCTTACTCAAGCGGCTGGTCTATCTCGAAGCTTCAGCTTACGACCT 1260  
 QY 1261 GCCAACGGCGAGTATGGCTACTCCGTGTGGAGTACTCGCGGCTGGGCTGA 1311  
 Db 1261 GCCAACGGCGAGTATGGCTACTCCGTGTGGAGTACTCGCGGCTGGGCTGA 1311

## RESULT 8

ABX08416  
 ID ABX08416 standard; DNA; 1311 BP.  
 AC ABX08416;  
 XX  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE DNA encoding alpha amylase polypeptide #6.  
 XX  
 KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX  
 OS Unidentified.

XX WO200268589-A2.  
 PN  
 XX  
 PD 06-SEP-2002.  
 XX  
 XX 21-FEB-2002; 2002WO-US005068.  
 PF  
 XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX  
 PA (DIVE-) DIVERSA CORP.

XX  
 PI Callen W, Richardson T, Frey G;  
 XX  
 XX P-PSDB; ABU03045.  
 DR WPI; 2003-018656/01.  
 DR P-PSDB; ABU03045.  
 XX  
 PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX  
 PS Claim 1; Fig 16; 30pp; English.  
 XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a

CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 XX

SQ Sequence 1311 BP; 321 A; 357 C; 382 G; 251 T; 0 U; 0 Other;

Query Match 89.4%; Score 1171.8; DB 7; Length 1311;  
 Best Local Similarity 93.4%; Pred. No. 0;  
 Matches 1224; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1 ATGCCCAAGTACTCCGAGCTGGAAAAGGGGGGTTCATATGCAGGCGTTCTACTCGGAC 60  
 Db 1 ATGCCCAAGTACTCCGAGCTGGAAAAGGGGGGTTCATATGCAGGCGTTCTACTCGGAC 60  
 QY 61 GTCCCTTACGAGGAATATGGTGGGACACAATACGGCAGAAAGATACCGAGTGGTACGAT 120  
 Db 61 GTCCCTTACGAGGAATATGGTGGGACACAATACGGCAGAAAGATACCGAGTGGTACGAT 120  
 QY 121 GCCGGATCTCCGCAATATGGATTCGCCGGGAGCAAGGGCATGGCGCGGCTTATTCG 180  
 Db 121 GCCGGATCTCCGCAATATGGATTCGCCGGGAGCAAGGGCATGGCGCGGCTTATTCG 180  
 QY 181 ATGGGCTACGACCCCTACGACTTCCTTTGACCTCGGTGAGTACGACCAAGGAAACGGTA 240  
 Db 181 ATGGGCTACGACCCCTACGACTTCCTTTGACCTCGGTGAGTACGACCAAGGAAACGGTA 240  
 QY 241 GAGACGCGCTTGGCTCCAGAGGAGCTCGTGAACATGATATAACACCGCCACGCTAT 300  
 Db 241 GAGACGCGCTTGGCTCCAGAGGAGCTCGTGAACATGATATAACACCGCCACGCTAT 300  
 QY 301 GGCATGAAGTATAGCCGATATAGTCATCAACACCGCGCGGTGACCTGGAAGTGG 360  
 Db 301 GGCATGAAGTATAGCCGATATAGTCATCAACACCGCGCGGTGACCTGGAAGTGG 360  
 QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCCTCAAGGTCGCGTGGGTAATAC 420  
 Db 361 AACCCCTTCGTGGGACTACACCTGGACCGACTTCCTCAAGGTCGCGTGGGTAATAT 420  
 QY 421 ACGGCCAAGTACTCGACTTCACCCGACGAGCTCCATGCGGGCGATTCCGGAACATTT 480  
 Db 421 ACTGCCAAGTACTCGACTTCACCCGACGAGTCAAGTGTGTGACAGGGGCAATTT 480  
 QY 481 GGAGGCTATCCCGACATATGCGCACCAAGAGCTGGGACCACTACTTGTGGGCGAGC 540  
 Db 481 GGAGGCTTCCGACATATAGCCCAAGAGAGCTGGGACCACTACTTGTGGGCGAGC 540  
 QY 541 CAGGAGAGTACCGGCAATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
 Db 541 GATGAGAGTACCGGCAATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
 QY 601 GTCAAGGGCTATGCTCCCTGGGTGCTCAAGGAGTGGCTCAACTGGTGGGAGGCTGGGG 660  
 Db 601 GTCAAGGGCTATGAGGAGTGGGTGCTCAAGGAGTGGCTCAACTGGTGGGAGGCTGGGG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATACTCGAGCGT 720  
 Db 661 GTCGGGGAGTACTGGGACACCAACGTTGATGCACTGCTCAACTGGGCCTACTCGAGCGAT 720  
 QY 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGATGAGGCGCTTTCACACCAAAAC 780  
 Db 721 GCMAAAGTCTTGAGCTTCCCGCTCTACTACAGATGATGAGGCGCTTTCACACCAAAAC 780



QY 781 ATTCCAGCGCTGCTCTCTCCCTTTCAGACGCGCAGACTGTTGTCTCCGCGACCCGTTT 840  
 Db |||||  
 QY 781 ATTCCAGCGCTGCTCTCTCCCTTTCAGACGCGCAGACTGTTGTCTCCGCGACCCGTTT 840  
 Db |||||  
 QY 841 AAGCGCGTAACTTTGTAGCAACACGACACCGATATAATCTGGAACAAGTATCCAGCC 900  
 Db |||||  
 QY 841 AAGCGCGTAACTTTGTAGCAACACGACACCGATATAATCTGGAACAAGTATCCAGCC 900  
 Db |||||  
 QY 901 TAGCGCTTATCTCTACCTACGAGCGCGCAGCAATATTCTACCGGACTACGAGGAG 960  
 Db |||||  
 QY 901 TAGCGCTTATCTCTACCTACGAGCGCGCAGCAATATTCTACCGGACTACGAGGAG 960  
 Db |||||  
 QY 961 TGCGTCAACAAGGATTAAGCTCAAGAACCTTCATCTGGATACATGAGAACTCTCCCGGAGGA 1020  
 Db |||||  
 QY 961 TGCGTCAACAAGGATTAAGCTCAAGAACCTTCATCTGGATACATGAGAACTCTCCCGGAGGA 1020  
 Db |||||  
 QY 1021 AGACCGCATAGTCTACTACGATACGATGATCACTCTCTGTCAGGAACGGTACGGG 1080  
 Db |||||  
 QY 1021 AGACCGCATAGTCTACTACGATACGATGATCACTCTCTGTCAGGAACGGTACGGG 1080  
 Db |||||  
 QY 1081 GACAACCGCGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGAGGTGGTT 1140  
 Db |||||  
 QY 1081 GACAACCGCGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGAGGTGGTT 1140  
 Db |||||  
 QY 1141 TATGTCGCGAAGTTCCGCGGCGGTGCATCCACGATATACCTGTTAACTCTGGAGGCTGG 1200  
 Db |||||  
 QY 1141 TATGTCGCGAAGTTCCGCGGCGGTGCATCCACGATATACCTGTTAACTCTGGAGGCTGG 1200  
 Db |||||  
 QY 1201 GTAGACAGTACTGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 Db |||||  
 QY 1201 GTAGACAGTACTGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 Db |||||  
 QY 1261 GCCAACGGCAGTATGGCTACTCCGTTGGAGCTACTGCGGGTGGGCTGA 1311  
 Db |||||  
 QY 1261 GCCAACGGCAGTATGGCTACTCCGTTGGAGCTACTGCGGGTGGGCTGA 1311  
 Db |||||

RESULT 9  
 ABX08440  
 ID ABX08440 standard; DNA; 1311 BP.  
 AC ABX08440;  
 XX  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE DNA encoding alpha amylase polypeptide #30.  
 DE  
 DE Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX  
 OS Unidentified.  
 OS  
 XX W020268589-A2.  
 PN  
 XX  
 XX 06-SEP-2002.  
 XX  
 XX 21-FEB-2002; 2002WO-US005068.  
 XX  
 XX 21-FEB-2001; 2001US-0270495P.  
 PR  
 XX 21-FEB-2001; 2001US-0270496P.  
 PR  
 XX 14-MAY-2001; 2001US-0291122P.  
 XX  
 XX (DIVE-) DIVERSA CORP.  
 PA  
 XX Callen W, Richardson T, Frey G;  
 XX  
 XX WPI; 2003-018656/01.  
 DR  
 XX P-PSDB; ABU03089.  
 DR

XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX  
 FS Claim 1; Fig 16; 301pp; English.  
 XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper,  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 XX  
 SQ Sequence 1311 BP; 324 A; 367 C; 377 G; 243 T; 0 U; 0 Other;  
 Query Match 88.5%; Score 1160.6; DB 7; Length 1311;  
 Best Local Similarity 92.8%; Pred. No. 0;  
 Matches 1217; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
 QY 1 ATGCGCAAGTACTCCGAGCTGGAAAGGGGGGGTCAATATCGAGGCTTCTACTGGAC 60  
 Db |||||  
 QY 1 ATGCGCAAGTACTCCGAGCTGGAAAGGGGGGGTCAATATCGAGGCTTCTACTGGAC 60  
 Db |||||  
 QY 61 GTGCTTCAGAGGATATGTTGGGACACAATACGCGAGAGATACCGGAGTGGTACGAT 120  
 Db |||||  
 QY 61 GTGCTTCAGAGGATATGTTGGGACACAATACGCGAGAGATACCGGAGTGGTACGAG 120  
 Db |||||  
 QY 121 GCGGAAATCTCCGCAATATGATTTCCCGCGGAGCAAGGATGAGCGCGCTATTCTG 180  
 Db |||||  
 QY 121 GCGGAAATCTCCGCAATATGATTTCCCGCGGAGCAAGGATGAGCGCGCTATTCTG 180  
 Db |||||  
 QY 181 ATGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAAACGTA 240  
 Db |||||  
 QY 181 ATGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAAACGTTG 240  
 Db |||||  
 QY 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACACCGCCACGCTAT 300  
 Db |||||  
 QY 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACACCGCCACGCTAT 300  
 Db |||||  
 QY 301 GGCATGAAGTATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTG 360  
 Db |||||  
 QY 301 GGCATGAAGTATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTG 360  
 Db |||||  
 QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTTCAAAGGTGCGCTCGGTAATAC 420  
 Db |||||  
 QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTTCAAAGGTGCGCTCGGTAATAC 420  
 Db |||||  
 QY 421 ACGGCCAACTACTCTGACTTCCACCGAGAGCTCCATCGGGCGATTCGGGAACATTT 480  
 Db |||||  
 QY 421 ACGGCCAACTACTCTGACTTCCACCGAGAGCTCCATCGGGCGATTCGGGAACATTT 480  
 Db |||||  
 QY 481 GGAGGCTATCCGACATATGCCACGAGAGTGGGACCAAGTACTGGCTCTCGGGCAGC 540  
 Db |||||  
 QY 481 GGAGGCTATCCGACATATGCCACGAGAGTGGGACCAAGTACTGGCTCTCGGGCAGC 540  
 Db |||||  
 QY 541 CAGGAGAGCTACCGGATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC 600  
 Db |||||  
 QY 541 CAGGAGAGCTACCGGATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC 600  
 Db |||||  
 QY 601 GTCAGGGCTATGCTCCCTGGTCTCAAGACTGGGCTGAATCTGGTGGGAGGCTGGCG 660  
 Db |||||



```

Db 601 GTCAGAGGCTACGAGCGTGGTCTGTCAGAGACTGGCTGGAGCTGGTGGAGGCTGGGCC 660
QY 661 GTTGAGAGTACTGGGACCAACAGTCGACGCTGTTCTCAATCGGCATATCGAGCGGT 720
Db 661 GTCGGGAGTACTGGGACCAACAGTCGACGCTGTTCTCAATCGGCATATCGAGCGGT 720
QY 721 GCCAAGGCTTTGACCTTCGCTCGCTCTACTACAGATGGATAGGCTTTGACACAAAC 780
Db 721 GCAAAAGCTTCGACTTCGCTCTACTACAGATGGATAGGCTTTGACACAAAC 780
QY 781 ATTCAGCGCTGCTCTGCTCGCTCTAGAACCGCCAGACTGTTGCTCCGCGACCGGTC 840
Db 781 ATTCAGCGCTGCTCTGCTCGCTCTAGAACCGCCAGACTGTTGCTCCGCGACCGGTC 840
QY 841 AAGCGCTTAACCTTTGTCAGCAACACGACACCGATATAATCTGGACAAAGATTCAGCC 900
Db 841 AAGCGCTTAACCTTTGTCAGCAACACGACACCGATATAATCTGGACAAAGATTCAGCC 900
QY 901 TACGGCTTCATCCTCACCTACGAGGCGCAGCGCAATATTCTACCGGACTACGAGGAG 960
Db 901 TACGGCTTCATCCTCACCTACGAGGCGCAGCGCAATATTCTACCGGACTACGAGGAG 960
QY 961 TGGCTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020
Db 961 TGGCTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020
QY 1021 AGCACCAGATAGTCTACTAGATAACGATGAATCTCTCTGTCAGGAACGGCTACGGG 1080
Db 1021 AGCACCAGATAGTCTACTAGATAACGATGAATCTCTCTGTCAGGAACGGCTACGGG 1080
QY 1081 GACAAAGCGGGCTTAACTCAATCACTACGAGCTGAGCAAGCGGAGGAGTGGGTT 1140
Db 1081 GACAAAGCGGGCTTAACTCAATCACTACGAGCTGAGCAAGCGGAGGAGTGGGTT 1140
QY 1141 TATGTGCGAAGTTCGCGGCGCTGCATCCACAGATATCTGTAACTCTGAGGCTGG 1200
Db 1141 TACGTTCGAAAGTTCGCGGAGCGTGCATCCACAGATATCTGTAACTCTGAGGCTGG 1200
QY 1201 GTAGACAAAGTACGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
Db 1201 GTGACAAAGTGGTGGACTCAAGCGGCTGGTGTACCTCGAGGCGCTTCCACGACCG 1260
QY 1261 GCCAAGCGGAGTATGGTACTCCGTGTGAGACTACTCGCGGGTGGGCTGA 1311
Db 1261 GCCAAGCGGAGTATGGTACTCCGTGTGAGACTACTCGCGGGTGGGCTGA 1311

```

## RESULT 10

ABX08437

ID ABX08437 standard; DNA; 1311 BP.

AC ABX08437;

DT 21-JAN-2003 (first entry)

DE DNA encoding alpha amylase polypeptide #27.

XX Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.

XX Unidentified.

XX WO200268589-A2.

XX 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005068.

PF

XX

PR 21-FEB-2001; 2001US-0270495P.

PR 21-FEB-2001; 2001US-0270496P.

PR 14-MAY-2001; 2001US-0291122P.

PA (DIVE-) DIVERSA CORP.

XX Callen W, Richardson T, Frey G;

DR WPI; 2003-018656/01.

DR P-PSDB; ABU03066.

XX

PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 liquefying starch, for textile desizing, for treating lignocellulosic  
 fibers, and for producing high-maltose or high-glucose syrup.

PS Claim 1; Fig 16; 30pp; English.

XX

CC The invention relates to a purified polypeptide with alpha-amylase  
 activity and the polynucleotide encoding it. The polypeptide is useful  
 for hydrolysing starch linkages, for catalysing the breakdown of a  
 starch, for modifying small molecules, for liquefying starch, for washing  
 an object, for textile desizing, for treating lignocellulosic fibers, for  
 improving fibre properties, for enzymatic de-linking of recycled paper  
 pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 polypeptide is also useful for preparing a dough or a baked product  
 prepared from the dough and in corn-wet milling processes, detergents,  
 baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 and starch modification in the paper and pulp industry, for removing  
 starch containing stains from a material and for reducing staling of  
 bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention

SQ Sequence 1311 BP; 308 A; 368 C; 379 G; 256 T; 0 U; 0 Other;

## Query Match

88.2%; Score 1155.8; DB 7; Length 1311;

Best Local Similarity 92.6%; Pred. No. 2e-314;

Matches 1214; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1 ATGCCCAAGTACTCCAGCTCGAAGAGCGCGGGTCAATATCGAGCGTTCCTACTGGGAC 60

Db 1 ATGCCCAAGTACTCCAGCTCGAAGAGCGCGGGTCAATATCGAGCGTTCCTACTGGGAC 60

QY 61 GTGCCCTTCAGGAGGAATATGTTGGGACACAATACGCGAGACATACCGAGTGTACGAT 120

Db 61 GTGCCCTTCAGGAGGAATATGTTGGGACACAATACGCGAGACATACCGAGTGTACGAT 120

QY 121 GCGGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCGCTATTTCG 180

Db 121 GCGGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCGCTATTTCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCTTGACCTCGGTGAGTACGACGAGGACGGTA 240

Db 181 ATGGGCTACGACCCCTACGACTTCTTGACCTCGGTGAGTACGACGAGGACGGTA 240

QY 241 GAGACGCGCTTTGGTCCCAAGCAGGAGCTCGTGAACATATATAACACCCGCCACGCTAT 300

Db 241 GAGACGCGCTTTGGTCCCAAGCAGGAGCTCGTGAACATATATAACACCCGCCACGCTAT 300

QY 301 GGCATGAAGGTAATAGCCGATATAGTATCAACACCGCGCGCGGCGGTGACCTGAGTGG 360

Db 301 GGCATGAAGGTAATAGCCGATATAGTATCAACACCGCGCGCGGCGGTGACCTGAGTGG 360

QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTCGCGGTAAATAC 420

Db 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTCGCGGTAAATAC 420

QY 421 ACGGCCAACTACTCGACTTCACCGGAAACGAGCTCCATTCGCGGCGATTCGCGGAACATTT 480

Db 421 ACGGCCAACTACTCGACTTCACCGGAAACGAGCTCCATTCGCGGCGATTCGCGGAACATTT 480

QY 481 GGAGGCTATCCCGACATATGCAACAGAGCTGGACCAAGTATGCTCTGGGCCAGC 540  
DB 481 CGAGGCTATCCCGACATATGCAACAGAGCTGGACCAAGTATGCTCTGGGCCAGC 540  
QY 541 CAGGAGCTAGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTGACTAC 600  
DB 541 CAGGAGCTAGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTGACTAC 600  
QY 601 GTCAAGGGCTATGCTCCCTGGCTCGTCAAGAGTGGCTGAACTGGTGGGAGGCTGGCG 660  
DB 601 GTCAAGGGCTATGCTCCCTGGCTCGTCAAGAGTGGCTGAACTGGTGGGAGGCTGGCG 660  
QY 661 GTTGAGAGTACTGGGACACCAAGCTGACGCTGTCTCAAGTGGGATCTCGAGCGGT 720  
DB 661 GTTGAGAGTACTGGGACACCAAGCTGACGCTGTCTCAAGTGGGATCTCGAGCGGT 720  
QY 721 GCCAAGGCTTTGACTTTCGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAAC 780  
DB 721 GCCAAGGCTTTGACTTTCGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAAC 780  
QY 781 ATTCCAGGCTCTCTCTGCTCTGAGACGGCCAGATGTTGTCTCCGCGACCCGTTT 840  
DB 781 ATTCCAGGCTCTCTCTGCTCTGAGACGGCCAGATGTTGTCTCCGCGACCCGTTT 840  
QY 841 AAGGCGTAACTTTGACCAACACGACACCGATATAATCTGGAAACAAGTATCCAGCC 900  
DB 841 AAGGCGTAACTTTGACCAACACGACACCGATATAATCTGGAAACAAGTATCCAGCC 900  
QY 901 TAGCGCTTATCTCTACCTACGAGGCGCCAGCCGACATATTTCTCCGAGTACGAGAG 960  
DB 901 TAGCGCTTATCTCTACCTACGAGGCGCCAGCCGACATATTTCTCCGAGTACGAGAG 960  
QY 961 TGGCTCAACAGATTAAGCTCAGAACCTCATCTGATATGATGAGTATGAGTATGAGT 1020  
DB 961 TGGCTCAACAGATTAAGCTCAGAACCTCATCTGATATGATGAGTATGAGTATGAGT 1020  
QY 1021 AGCAGCGATAGTCTTACTACGATACGATGAATCATCTCTCGTCAAGAACGGTACGGG 1080  
DB 1021 AGCAGCGATAGTCTTACTACGATACGATGAATCATCTCTCGTCAAGAACGGTACGGG 1080  
QY 1081 GACAACGGGGTTATACCTATCATCACTAGCTAGCTCAGCAAGCGGAGGTTGGTT 1140  
DB 1081 AGCAAGCTGGCTTATATCTTACATCACTCGCTCGGCTCAGCAAGGTTGGAGGTTGGTT 1140  
QY 1141 TATGTCGCAAGTTCGCGGCGCTGCATCCAGAGTATGATGATGATGATGATGATGATG 1200  
DB 1141 TATGTCGCAAGTTCGCGGCGCTGCATCCAGAGTATGATGATGATGATGATGATGATG 1200  
QY 1201 GTAGACAAGTACTTACTCAAGCGCTGGTCTATCTCGAAGTCTCAGCTTACGACCT 1260  
DB 1201 GTAGACAAGTACTTACTCAAGCGCTGGTCTATCTCGAAGTCTCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGCTATGCTCTCTCTGAGTCTGAGTCTGAGGCTGAGGCTGAGGCTGAG 1311  
DB 1261 GCCAAGCGGCTATGCTCTCTCTGAGTCTGAGTCTGAGGCTGAGGCTGAGGCTGAG 1311

RESULT 11  
ID ABX08422  
XX ABX08422 standard; DNA; 1311 BP.  
AC ABX08422;  
DT 21-JAN-2003 (first entry)  
XX  
DE DNA encoding alpha amylase polypeptide #12.  
XX Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW

starch modification.  
Unidentified.  
W0200268589-A2.  
06-SEP-2002.  
21-FEB-2002; 2002WO-US005068.  
21-FEB-2001; 2001US-0270495P.  
14-FEB-2001; 2001US-0270496P.  
14-MAY-2001; 2001US-0291122P.  
(DIVE-) DIVERSA CORP.  
Callen W, Richardson T, Frey G;  
WPI; 2003-018656/01.  
P-PSDB; ABU03051.  
Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
liquefying starch, for textile desizing, for treating lignocellulosic  
fibers, and for producing high-maltose or high-glucose syrup.  
Claim 1; Fig 16; 301pp; English.  
The invention relates to a purified polypeptide with alpha-amylase  
activity and the polynucleotide encoding it. The polypeptide is useful  
for hydrolysing starch linkages, for catalysing the breakdown of a  
starch, for modifying small molecules, for liquefying starch, for washing  
an object, for textile desizing, for treating lignocellulosic fibers, for  
improving fibre properties, for enzymatic de-inking of recycled paper  
pulp, for producing a high-maltose or high-glucose syrup or a mixed  
syrup, and for increasing the flow of production fluids from a  
subterranean formation by removing a viscous, starch-containing, damaging  
fluid formed during production operations and found within the  
subterranean formation which surrounds a completed well bore. The  
polypeptide is also useful for preparing a dough or a baked product  
prepared from the dough and in corn-wet milling processes, detergents,  
baking processes, beverages, oil fields (fuel ethanol), brewing processes  
and starch modification in the paper and pulp industry, for removing  
starch containing stains from a material and for reducing staling of  
bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
encoding alpha amylase polypeptides of the invention  
Sequence 1311 BP; 313 A; 369 C; 381 G; 248 T; 0 U; 0 Other;

Query Match 87.68; Score 1147.8; DB 7; Length 1311;  
Best Local Similarity 92.28; Pred. No. 3e-312;  
Matches 1209; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
QY 1 ATGCCCAAGTACTCCGAGCTCGAAGAGGGCGGGTCTATATGAGCGGTTCTACTGGAC 60  
DB 1 ATGCCCAAGTACTCCGAGCTCGAAGAGGGCGGGTCTATATGAGCGGTTCTACTGGAC 60  
QY 61 GTGCTTCAGAGGATATGTTGGGACACATACGGCAGAGATACCGGAGGTACGAT 120  
DB 61 GTGCTTCAGAGGATATGTTGGGACACATACGGCAGAGATACCGGAGGTACGAT 120  
QY 121 GCCCGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGGCATGGCGCGCTATTTCG 180  
DB 121 GCCCGAATCTCCGCGATATGATTTCCCGCGGAGCAAGGGCATGGCGCGCTATTTCG 180  
QY 181 ATGGGCTACAGCCCTTACGACTCTTTGACTCGGTGAGTACCGACAGAGGGAACGTA 240  
DB 181 ATGGGCTACAGCCCTTACGACTCTTTGACTCGGTGAGTACCGACAGAGGGAACGTA 240  
QY 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACCCGCCACGCTAT 300  
DB 241 GAAACAGGTTTCGGCTCAAGAGCAGGAGCTCATAAACATGATAAACCCGCCACGCTAT 300  
QY 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACCCGCCCGCGGTGACCTGGAGTGG 360

RESULT 12  
ABX08432  
ID ABX0  
XX  
AC ABX0

XX	21-JAN-2003	(first entry)
XX		
XX	DNA encoding alpha amylase polypeptide #22.	
XX		
KW	Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;	
KW	starch breakdown catalysis; textile desizing; lignocellulosic fibre;	
KW	enzymatic de-linking; recycled paper; high-maltose syrup; dough;	
KW	high glucose syrup; corn-wet milling process; detergent; baking process;	
KW	beverage; oil field; fuel ethanol; brewing process; staling;	
KW	starch modification.	
XX		
OS	Unidentified.	
XX		
XX	WO200268589-A2.	
XX		
XX	06-SEP-2002.	
XX		
XX	21-FEB-2002; 2002WO-US0005068.	
XX		
PR	21-FEB-2001; 2001US-0270495P.	
PR	21-FEB-2001; 2001US-0270496P.	
PR	14-MAY-2001; 2001US-0291122P.	
XX		
XX	(DIVE-) DIVERSA CORP.	
XX		
XX	Callen W, Richardson T, Frey G;	
XX		
DR	WPI; 2003-018656/01.	
DR	P-PSDB; ABU03061.	
XX		
PT	Novel purified polypeptide with alpha-amylase activity, useful e.g. for	
PT	liquefying starch, for textile desizing, for treating lignocellulosic	
PT	fibers, and for producing high-maltose or high-glucose syrup.	
XX		
XX	Claim 1; Fig 16; 301pp; English.	
XX		
CC	The invention relates to a purified polypeptide with alpha-amylase	
CC	activity and the polynucleotide encoding it. The polypeptide is useful	
CC	for hydrolysing starch linkages, for catalysing the breakdown of a	
CC	starch, for modifying small molecules, for liquefying starch, for washing	
CC	an object, for textile desizing, for treating lignocellulosic fibers, for	
CC	improving fibre properties, for enzymatic de-linking of recycled paper	
CC	pulp, for producing a high-maltose or high-glucose syrup or a mixed	
CC	syrup, and for increasing the flow of production fluids from a	
CC	subterranean formation by removing a viscous, starch-containing, damaging	
CC	fluid formed during production operations and found within the	
CC	subterranean formation which surrounds a completed well bore. The	
CC	polypeptide is also useful for preparing a dough or a baked product	
CC	prepared from the dough and in corn-wet milling processes, detergents,	
CC	baking processes, beverages, oil fields (fuel ethanol), brewing processes	
CC	and starch modification in the paper and pulp industry, for removing	
CC	starch containing stains from a material and for reducing staling of	
CC	bakery products. Sequences ABX08411-ABX08515 represent DNA molecules	
CC	encoding alpha amylase polypeptides of the invention	
XX		
SQ	Sequence 1311 BP; 313 A; 365 C; 382 G; 251 T; 0 U; 0 Other;	
	Query Match 87.6%; Score 1147.8; DB 7; Length 1311;	
	Best Local Similarity 92.2%; Pred. No. 3e-312;	
	Matches 1209; Conservative 0; Mismatches 102; Indels 0; Gaps 0;	
Qy	1 ATGCCAAGTACTCCGAGCTGGAAAAGGCGGGGTGTAATCGAGCGTCTTACTGGGAC 60	
Db	1 ATGCCAAGTACTCCGACCTGGAAGCGGCGGTATATAATCGAGGCTCTTACTGGGAC 60	
Qy	61 GTGCCCTTCAGGAGGATATGTTGGGACACAATACGCAGCAATACCGAGTGGTACGAT 120	
Db	61 GTCCAGGTGGAGGATCTGGTGGGACACATCAGGAGCAAGATACCGAGTGGTACGAG 120	
Qy	121 GCGGGAATCTCCGCAATATGGAATCTCCCGGGGAGCAAGGGGATGGGGCGGCGCTATTGC 180	
Db	121 GCGGGAATCTCCGCAATTTGGAATCTCCCGGGGAGCAAGGGGATGGGGCGGCGCTATTGC 180	

QY 181 ATGGCTAGGACCTTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA 240  
 Db 181 ATGGCTAGGACCTTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA 240  
 QY 241 GAGAGCGCTTTGGCTTCCAAAGCAGAGCTCGTGAACATGATAAACAACCGCCACGCTTAT 300  
 Db 241 GAGAGCGCTTTGGCTTCCAAAGCAGAGCTCGTGAACATGATAAACAACCGCCACGCTTAT 300  
 QY 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACACCGCGCGCGGTGAGCTGAGTGG 360  
 Db 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACACCGCGCGCGGTGAGCTGAGTGG 360  
 QY 361 AACCCTTCTGTAAGACTATACCTTGGACCGACTTCTCAAGAGTCCGCTCGGTGAATAATAC 420  
 Db 361 AACCCTTCTGTAAGACTATACCTTGGACCGACTTCTCAAGAGTCCGCTCGGTGAATAATAC 420  
 QY 421 ACGGCAACTACTCGACTTCCACCGAAGAGCTCCATGCGGGGATTCGGGAACATTT 480  
 Db 421 ACGGCAACTACTCGACTTCCACCGAAGAGCTCCATGCGGGGATTCGGGAACATTT 480  
 QY 481 GGAGGCTATCCGACATATGCCACGACAGAGCTGCGACCTACTGCTCTGGGCCAGC 540  
 Db 481 GGAGGCTATCCGACATATGCCACGACAGAGCTGCGACCTACTGCTCTGGGCCAGC 540  
 QY 541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATGCTGCTGCGCTTCGACTTAC 600  
 Db 541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATGCTGCTGCGCTTCGACTTAC 600  
 QY 601 GTCAAGGGCTATGCTCCCTGGTGTCTCAAGGACTGGCTGACTGCTGGGAGGCTGGCG 660  
 Db 601 GTCAAGGGCTATGCTCCCTGGTGTCTCAAGGACTGGCTGACTGCTGGGAGGCTGGCG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACCGAGCTGCTTCTCAACTGGGCATCTCGAGCGGT 720  
 Db 661 GTTGGAGAGTACTGGGACACCAACCGAGCTGCTTCTCAACTGGGCATCTCGAGCGGT 720  
 QY 721 GCAAGGCTTTGACTTGGCTTCTACTCAAGATGATGAGGCTTTGACACCAAAAC 780  
 Db 721 GCAAGGCTTTGACTTGGCTTCTACTCAAGATGATGAGGCTTTGACACCAAAAC 780  
 QY 781 ATTCCAGCGCTGCTCTGCTTCTAGACGCGAGCTGTTGTCTCCGCGACCGCTTC 840  
 Db 781 ATTCCAGCGCTGCTCTGCTTCTAGACGCGAGCTGTTGTCTCCGCGACCGCTTC 840  
 QY 841 AAGCGCGTAACTTTGTAGCAACACGACACCGATATATCTGGAACAAGTATCCAGCC 900  
 Db 841 AAGCGCGTAACTTTGTAGCAACACGACACCGATATATCTGGAACAAGTATCCAGCC 900  
 QY 901 TAGCGCTTATCTCTACCTACGAGGCGGCGGACGACATATCTACCGGACTACGAGAG 960  
 Db 901 TAGCGCTTATCTCTACCTACGAGGCGGCGGACGACATATCTACCGGACTACGAGAG 960  
 QY 961 TGCGTCAACAAGGATTAAGTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGA 1020  
 Db 961 TGCGTCAACAAGGATTAAGTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGA 1020  
 QY 1021 AGCAGGACATAGTCTACTAGTACGATGAAGTCACTCTTCTGTCAGGAACGCTACGGG 1080  
 Db 1021 AGCAGGACATAGTCTACTAGTACGATGAAGTCACTCTTCTGTCAGGAACGCTACGGG 1080  
 QY 1081 GACAAGCGGGGCTTAACTCACTCAACCTAGGCTCGAGCAAGCGCCGGAAGGTTGGT 1140  
 Db 1081 GACAAGCGGGGCTTAACTCACTCAACCTAGGCTCGAGCAAGCGCCGGAAGGTTGGT 1140  
 QY 1141 TATGTGCGGAAGTTGCGGGCGGTGATCCAGAGTATCTGTTAACTTCGAGGCTGG 1200  
 Db 1141 TATGTGCGGAAGTTGCGGGCGGTGATCCAGAGTATCTGTTAACTTCGAGGCTGG 1200  
 QY 1201 GTAGACAGTACGCTTACTCAAGCGGTGGTCTTCTGAGCTCCAGCTTACGACCT 1260  
 Db 1201 GTAGACAGTACGCTTACTCAAGCGGTGGTCTTCTGAGCTCCAGCTTACGACCT 1260

QY 1261 GCCAACGGCAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGCTGA 1311  
 Db 1261 GCCAACGGCAGTATGGCTACTCCGTGTGGAGTATTCGGGTGTGGCTGA 1311  
 RESULT 13  
 ABX08412  
 ID ABX08412 standard; DNA; 1311 BP.  
 AC ABX08412;  
 XX 21-JAN-2003 (first entry)  
 XX DNA encoding alpha amylase polypeptide #2.  
 DE Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 XX starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX Unidentified.  
 OS W0200268589-A2.  
 XX 06-SEP-2002.  
 XX 21-FEB-2002; 2002WO-US005068.  
 XX 21-FEB-2001; 2001US-0270495P.  
 XX 21-FEB-2001; 2001US-0270496P.  
 XX 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Callen W, Richardson T, Frey G;  
 WPI; 2003-018656/01.  
 P-PSDB; ABU03041.  
 Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 liquefying starch, for textile desizing, for treating lignocellulosic  
 fibers, and for producing high-maltose or high-glucose syrup.  
 Claim 1; Fig 16; 301pp; English.  
 The invention relates to a purified polypeptide with alpha-amylase  
 activity and the polynucleotide encoding it. The polypeptide is useful  
 for hydrolysing starch linkages, for catalysing the breakdown of a  
 starch, for modifying small molecules, for liquefying starch, for washing  
 an object, for textile desizing, for treating lignocellulosic fibers, for  
 improving fibre properties, for enzymatic de-inking of recycled paper  
 pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 syrup, and for increasing the flow of production fluids from a  
 subterranean formation by removing a viscous, starch-containing  
 fluid formed during production operations and found within the  
 subterranean formation which surrounds a completed well bore. The  
 polypeptide is also useful for preparing a dough or a baked product  
 prepared from the dough and in corn-wet milling processes, detergents,  
 baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 and starch modification in the paper and pulp industry, for removing  
 starch containing stains from a material and for reducing staling of  
 bakery products. Sequences ABX08411-ABX08415 represent DNA molecules  
 encoding alpha amylase polypeptides of the invention  
 SQ Sequence 1311 BP; 313 A; 378 C; 376 G; 244 T; 0 U; 0 Other;  
 Query Match 87.4%; Score 1146.2; DB 7; Length 1311;  
 Best Local Similarity 92.1%; Pred. No. 8.5e-312;  
 Matches 1208; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
 QY 1 ATGGCCAGTACTCCGAGCTGGAAAGCGGGGTGTCATATGCGAGCGTTCTACTGGGAC 60

Db 1 ATGSCCAAGTACCTGGAGCTGGAAGAGGGCGGGCTCATATGTCAGGCTTCTACTGGGAC 60  
QY 61 GTGCTTCAGGAGGATATGTGGGACACATATCGGAGAGATACCGGAGTGTACGAT 120  
Db 61 GTCCCATGGAGGATCTGTGGGACACGATAGCCAGAGATATACCCGACTGGGCAAGC 120  
QY 121 GCCGGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGSCATGGGCGGCGCTATTTCG 180  
Db 121 GCCGGGATTTCCGCGATATGATTTCCCGCGGAGCAAGGSCATGGGCGGCGCTATTTCG 180  
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGTGAGTACGACAGAGGAGAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGTGAGTACGACAGAGGAGAACGGTA 240  
QY 241 GAGACGGCTTTGGCTCAAGCAGGAGCTTCGTGAACATGATAAACAACCGCCACCGCTAT 300  
Db 241 GAGACGGCTTTGGCTCAAGCAGGAGCTTCGTGAACATGATAAACAACCGCCACCGCTAT 300  
QY 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGAGATGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGAGATGG 360  
QY 361 AACCCCTTCGTGAAGGACTATACCTGGAGCGGACTTCTCAAGGTCGGTGGGTAATATAC 420  
Db 361 AACCCCTTCGTGAAGGACTATACCTGGAGCGGACTTCTCAAGGTCGGTGGGTAATATAC 420  
QY 421 ACGGCCAATACCTCGACTTCCACCCGAGACGAGTCCATCGGCGGATTCGGAAACATTT 480  
Db 421 ACGGCCAATACCTCGACTTCCACCCGAGACGAGTCCATCGGCGGATTCGGAAACATTT 480  
QY 481 GGAGGCTATCCGACATATGCGACGACGACGAGCTGGACCAAGTACTGGCTTGGGCCAGC 540  
Db 481 GGAGGCTATCCGACATATGCGACGACGACGAGCTGGACCAAGTACTGGCTTGGGCCAGC 540  
QY 541 CAGGAGAGCTACGGGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGTTCGACTAC 600  
Db 541 CAGGAGAGCTACGGGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGTTCGACTAC 600  
QY 601 GTCAGGGCTATGCTCCCTGGTGGTTCGAGGCTGCTGAGGCTGCTGGGAGGCTGGGCG 660  
Db 601 GTCAGGGCTATGCTCCCTGGTGGTTCGAGGCTGCTGAGGCTGCTGGGAGGCTGGGCG 660  
QY 661 GTTGAGAGTACTGGGACCAACAGTCGACGCTGCTTCTCAACTGGGATCTCGAGCGGT 720  
Db 661 GTTGAGAGTACTGGGACCAACAGTCGACGCTGCTTCTCAACTGGGATCTCGAGCGGT 720  
QY 721 GCCAAGGCTTTTGAATTCGCTTCTACTACAAGATGGATGAGGCTTTGACAAACAAAAC 780  
Db 721 GCCAAGGCTTTTGAATTCGCTTCTACTACAAGATGGATGAGGCTTTGACAAACAAAAC 780  
QY 781 ATTCAGGCTGCTGCTGCTGCTTCAAGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 ATTCAGGCTGCTGCTGCTGCTTCAAGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 AAGGCGCTTAAGCTTTGTAGCAAAACACGACACCGATATAATCTGGAACAGTATCCAGCC 900  
Db 841 AAGGCGCTTAAGCTTTGTAGCAAAACACGACACCGATATAATCTGGAACAGTATCCAGCC 900  
QY 901 TAGGGCTTATCTCCTACCTAGAGGCGGAGGACATATTTCTACCGGACTACGAGGAG 960  
Db 901 TAGGGCTTATCTCCTACCTAGAGGCGGAGGACATATTTCTACCGGACTACGAGGAG 960  
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGCGGAGGA 1020  
QY 1021 AGCACCGCATAGTCTACTACGATAACGATGAATTCATCTTCGAGAAACGGTACGGG 1080  
Db 1021 AGCACTGACATCGTTTACTAGCAACACGACGAGTGTATTCGTGAGAAACGGTACGGG 1080  
QY 1081 GACAGCGGGGCTTAACTACATCACTAGCTCGACAGCGCGGAGTGGGT 1140

Db 1081 AGCAAGCCGGGACTGATACATACATACATCACTCGCTCAAGCGAAGCCGGAAGTGGTTC 1140  
QY 1141 TATGTCCCAAGTTCGCGGCGCTGATCCACAGAGTATACCTGTAACCTCGAGGCTGG 1200  
Db 1141 TACGTTCCCAAGTTCGCGGCGCTGATCCACAGAGTATACCTGTAACCTCGAGGCTGG 1200  
QY 1201 GTAGACAAGTACGCTTACTCAAGCGGCTGGGTATCTCGAAGCTCCAGCTTACGACCTT 1260  
Db 1201 GTGACAAGTGGGTGGGTGACTCAAGCGGCTGGGTGATCTCGAGGCGCTGCCACGACCGG 1260  
QY 1261 GCCAAGCGGCACTATGGCTACTCCGCTGGAGCTACTCGGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGGCTATACGGCTACTCCGCTGGAGCTATTCGGGTGGGCTGA 1311

RESULT 14

ABX08439  
ID ABX08439 standard; DNA; 1311 BP.  
XX ABX08439;  
XX AC  
XX DT 21-JAN-2003 (first entry)  
XX XX  
DE DNA encoding alpha amylase polypeptide #29.  
XX Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.

XX Unidentified.  
XX OS  
XX PN WO200268589-A2.  
XX PD 06-SEP-2002.  
XX PF 21-FEB-2002; 2002WO-US005068.  
XX PR 21-FEB-2001; 2001US-0270495P.  
XX PR 21-FEB-2001; 2001US-0270496P.  
XX PR 14-MAY-2001; 2001US-0291122P.  
XX PA (DIVE-) DIVERSA CORP.  
XX PI Callen W, Richardson T, Frey G;  
XX DR WPI; 2003-018656/01.  
XX DR P-PSDB; ABU03068.  
XX PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
PT liquefying starch, for textile desizing, for treating lignocellulosic  
PT fibers, and for producing high-maltose or high-glucose syrup.

Claim 1; Fig 16; 301pp; English.

CC The invention relates to a purified polypeptide with alpha-amylase  
CC activity and the polynucleotide encoding it. The polypeptide is useful  
CC for hydrolysing starch linkages, for catalysing the breakdown of a  
CC starch, for modifying small molecules, for liquefying starch, for washing  
CC an object, for textile desizing, for treating lignocellulosic fibers, for  
CC improving fibre properties, for enzymatic de-linking of recycled paper  
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of

CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention

XX SQ Sequence 1311 BP; 321 A; 361 C; 377 G; 252 T; 0 U; 0 Other;

Query Match 87.4%; Score 1146.2; DB 7; Length 1311;  
 Best Local Similarity 92.1%; Pred. No. 8.5e-312;  
 Matches 1208; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 1 ATGGCCAACTCTCCAGCTCGAAGTGGGAGGCGGGTCAATATCCAGCGGTTCCTACTGGGAC 60  
 DB 1 ATGGCCAACTCTCCAGCTCGAAGTGGGAGGCGGGTCAATATCCAGCGGTTCCTACTGGGAC 60

QY 61 GTGCGCTTCAGGAGGAATATGGTGGGACAAATACCGCAGAAATACCGGAGTGGTACGAT 120  
 DB 61 GTGCGCTTCAGGAGGAATATGGTGGGACAAATACCGCAGAAATACCGGAGTGGTACGAT 120

QY 121 GCGGGAATCTCCGCAATATGGATTCGCCCGGAGCAAGGCGATGGCGCGCCCTATTTCG 180  
 DB 121 GCGGGAATCTCCGCAATATGGATTCGCCCGGAGCAAGGCGATGGCGCGCCCTATTTCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCCTTGACCTCGGTGAGTACGACCAAGGAGCAAGGTA 240  
 DB 181 ATGGGCTACGACCCCTACGACTTCCTTGACCTCGGTGAGTACGACCAAGGAGCAAGGTA 240

QY 241 GAGACCGCTTTGGCTCCAGCAGGAGCTCGTGAACATGATATAAACACCGCCACGCTAT 300  
 DB 241 GAAACGAGGTTCGGCTCAAGCAGGAGCTCATATAACATGATAAACACCGCCACGCTAT 300

QY 301 GGCATGAAGGTATAGCCATATAGTATCATACCAACCGCGCGGCGGTGACCTGGAGTGG 360  
 DB 301 GGCATGAAGGTATAGCCATATAGTATCATACCAACCGCGCGGCGGTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTGAACGACTATACCTGGACGAGTCTCAAGAGTCGCGTGGGTAATAC 420  
 DB 361 AACCCCTTCGTGAACGACTATACCTGGACGAGTCTCAAGAGTCGCGTGGGTAATAC 420

QY 421 ACGGCCAACTACTCTGACTTCCACCGGAACGAGTCCATGCGGGGAGATTCGGGAACATTT 480  
 DB 421 ACGGCCAACTACTCTGACTTCCACCGGAACGAGTCCATGCGGGGAGATTCGGGAACATTT 480

QY 481 GGAGGCTATCCGACATATGCGACGACAGAGCTGGACCACTACTGCTCTGGGCCAGC 540  
 DB 481 GGAGGCTATCCGACATATGCGACGACAGAGCTGGACCACTACTGCTCTGGGCCAGC 540

QY 541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
 DB 541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600

QY 601 GTCAAGGGCTATGCTCCCTGGGTCTCAGGACTGGCTGAACCTGGTGGGAGGCTGGCG 660  
 DB 601 GTCAAGGGCTATGCTCCCTGGGTCTCAGGACTGGCTGAACCTGGTGGGAGGCTGGCG 660

QY 661 GTTGAGAGTACTGGGACACCAACGCTCGACGCTGTTCTCAACTGGGCATACTCGAGCGGT 720  
 DB 661 GTTGAGAGTACTGGGACACCAACGCTCGACGCTGTTCTCAACTGGGCATACTCGAGCGGT 720

QY 721 GCGAAGTCTTTGACTTCGCTCCCTACTACAGATGATGAGGCTTTGACACAAAC 780  
 DB 721 GCGAAGTCTTTGACTTCGCTCCCTACTACAGATGATGAGGCTTTGATAAACAAAC 780

QY 781 ATTCCAGCGCTCTGCTCTCCCTTCAGAACCGGCGAGCTGTTGTCTCCCGGACCCGCTTC 840  
 DB 781 ATTCCAGCGCTCTGCTCTCCCTTCAGAACCGGCGAGCTGTTGTCTCCCGGACCCGCTTC 840

QY 841 AGGCGGTAACCTTTGTAGAACCAACGACCGATATATCTGGAAACAGTATCCAGCC 900  
 DB 841 AGGCGGTAACCTTTGTAGAACCAACGACCGATATATCTGGAAACAGTATCCAGCC 900

QY 901 TAGCGGTTCTATCTCCTACCTACGAGGCGGCGGACAAATATTTCTACCGGACTACGAGGAG 960  
 DB 901 TAGCGGTTCTATCTCCTACCTACGAGGCGGCGGACAAATATTTCTACCGGACTACGAGGAG 960

QY 961 TGGCTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGGAGA 1020  
 DB 961 TGGCTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGGAGA 1020

QY 1021 AGCACCGACATAGTCTACTACGATAACGATGAATCACTTCTGTCAGGAACGGCTACGGG 1080  
 DB 1021 AGCACTGACATCTGTTTACTACGACAACGACGAGCTGATATTCGTGAGAAACGGCTACGGA 1080

QY 1081 GACAGCCCGGGCTTAACTACATCAACCTAGGCTCGAGCAAGCCGGAAGTGGTT 1140  
 DB 1081 AGAAGCCGGGACTGATACATACATCAACCTCGCTCAAGCAAGCCGGAAGTGGTT 1140

QY 1141 TATGTGCGGAAGTTCGCGGCGGCTGCATCCACGAGTACTTGGTAACTTCGAGGCTGG 1200  
 DB 1141 TATGTGCGGAAGTTCGCGGCGGCTGCATCCACGAGTACTTGGTAACTTCGAGGCTGG 1200

QY 1201 GTAGACAAAGTACGCTTACTCAAGCGGCTGGGTCTATCTGMAAGCTCGAGCTTACGACCT 1260  
 DB 1201 GTAGACAAAGTACGCTTACTCAAGCGGCTGGGTCTATCTGMAAGCTCGAGCTTACGACCT 1260

QY 1261 GCCAAAGCGGCGAGTATGGCTACTCCGCTGTGGAGCTACTGCGGGGTGGGCTCA 1311  
 DB 1261 GCCAAAGCGGCGAGTATGGCTACTCCGCTGTGGAGCTACTGCGGGGTGGGCTCA 1311

## RESULT 15

ABX08435  
 ID ABX08435 standard; DNA; 1311 BP.

XX AC ABX08435;

XX AC ABX08435;

XX DT 21-JAN-2003 (first entry)

XX DE DNA encoding alpha amylase polypeptide #25.

XX KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 enzymatic de-inking; recycled paper; high-maltose syrup; dough; baking process;  
 high glucose syrup; corn-wet milling process; detergent; brewing process;  
 beverage; oil field; fuel ethanol; starch modification.

XX OS Unidentified.

XX FN WO200268589-A2.

XX PD 06-SEP-2002.

XX PF 21-FEB-2002; 2002WO-US005068.

XX PR 21-FEB-2001; 2001US-0270495P.

XX PR 21-FEB-2001; 2001US-0270496P.

XX PR 14-MAY-2001; 2001US-0291122P.

XX PA (DIVE-) DIVERSA CORP.

XX PI Callen W, Richardson T, Frey G;

XX XX WPI; 2003-018656/01.

XX DR P-PSDB; ABU03064.

XX XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 liquefying starch, for textile desizing, for treating lignocellulosic  
 fibers, and for producing high-maltose or high-glucose syrup.

XX PS Claim 1; Fig 16; 301pp; English.

XX XX The invention relates to a purified polypeptide with alpha-amylase  
 activity and the polynucleotide encoding it. The polypeptide is useful  
 for hydrolysing starch linkages, for catalysing the breakdown of a  
 starch, for modifying small molecules, for liquefying starch, for washing  
 an object, for textile desizing, for treating lignocellulosic fibers, for  
 improving fibre properties, for enzymatic de-inking of recycled paper

CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
XX encoding alpha amylase polypeptides of the invention  
SQ Sequence 1311 BP; 317 A; 371 C; 378 G; 245 T; 0 U; 0 Other;

Query Match 87.3%; Score 1144.6; DB 7; Length 1311;  
Best Local Similarity 92.1%; P: 0; Mismatches 104; Indels 0; Gaps 0;  
Matches 1207; Conservative 0;

QY 1 ATGGCCAAAGTACTCGAGCTCGAAGAGCGCGGGTCAATATGACAGGCGTCTTACTGGGAC 60  
DB 1 ATGGCCAAAGTACTCGAGCTCGAAGAGCGCGGGTCAATATGACAGGCGTCTTACTGGGAC 60  
QY 61 GTGGCTTCAGGAGGAATATGTTGGGACACAATAGCGGAGAGATACCGGAGTGTAGCAT 120  
DB 61 GTGGCTTCAGGAGGAATATGTTGGGACACAATAGCGGAGAGATACCGGAGTGTAGCAT 120  
QY 121 GCCGGAATCTCGCAATATGATATCCCGCGGAGCAGGCGATGGCGGCGCTATTTCG 180  
DB 121 GCCGGAATCTCGCAATATGATATCCCGCGGAGCAGGCGATGGCGGCGCTATTTCG 180  
QY 181 ATGGCTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
DB 181 ATGGCTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
QY 241 GAGAGCGCTTTGGCTCAAGCAGGAGCTCGTGAACATGATGAATGAACCGGCCACGCTAT 300  
DB 241 GAGAGCGCTTTGGCTCAAGCAGGAGCTCGTGAACATGATGAATGAACCGGCCACGCTAT 300  
QY 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
DB 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
QY 361 AACCCCTTCGTGAAGCACTATACCTGGACCGACTTCTCAAAAGTTCGGTTCGGGTAATAC 420  
DB 361 AACCCCTTCGTGAAGCACTATACCTGGACCGACTTCTCAAAAGTTCGGTTCGGGTAATAC 420  
QY 421 ACGGCAACTACCTCGACTTCCACCGGAACGAGCTCCATGCGGCGATTCCGGGAACATTT 480  
DB 421 ACGGCAACTACCTCGACTTCCACCGGAACGAGCTCCATGCGGCGATTCCGGGAACATTT 480  
QY 481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGACCAAGTACTGGCTCTGGGCCAGC 540  
DB 481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGACCAAGTACTGGCTCTGGGCCAGC 540  
QY 541 CAGGAGAGTACCGGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGTTGACTAC 600  
DB 541 GATGAGAGTACCGGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGTTGACTAC 600  
QY 601 GTCAAGGGCTATGCTCCCTGGGTCTGTCAGGAGTGGCTGAACTGGTGGGAGGCTGGGCG 660  
DB 601 GTGAGGGCTACGAGCGTGGGTCTGTCAGGAGTGGCTGAACTGGTGGGAGGCTGGGCG 660  
QY 661 GTTGAGAGTACTGGGACACCAACGTGACGCTGTTCTCAACTGGGCGATCTCGAGCGGT 720  
DB 661 GTTGAGAGTACTGGGACACCAACGTGACGCTGTTCTCAACTGGGCGATCTCGAGCGGT 720  
QY 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGGATGAGGCTTTGACACAAAC 780  
DB 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGGATGAGGCTTTGACACAAAC 780  
QY 781 ATTCAGCGCTCGTCTCTGCCCTTCAGAACCGGCCAGACTGTGTCTCCCGGACCGGTTTC 840

DB 781 ATTCCAGCGCTCGTCTCTGCCCTTCAGAACCGGCCAGACTGTGTCTCCCGGACCGGTTTC 840  
QY 841 AAGGCCGTAACTTTGTAGCAAAACACGACACCGATATTAATCTGGAACAAGTATCCAGCC 900  
DB 841 AAGGCCGTAACTTTGTAGCAAAACACGACACCGATATTAATCTGGAACAAGTATCCAGCC 900  
QY 901 TAGCGCTTCATCTCTACCTACGAGGCGCAGCGACAATATTCTACCGGCACTACGAGGAG 960  
DB 901 TAGCGCTTCATCTCTACCTACGAGGCGCAGCGACAATATTCTACCGGCACTACGAGGAG 960  
QY 961 TGCTCTCAACAAGGATAAGCTCAAGAACCTTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020  
DB 961 TGCTCTCAACAAGGATAAGCTCAAGAACCTTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020  
QY 1021 AGCACCGACATAGTCTACTACGATACGATGAACTCATCTTCGTGAGGAACGCTACGG 1080  
DB 1021 AGCACCGACATAGTCTACTACGATACGATGAACTCATCTTCGTGAGGAACGCTACGG 1080  
QY 1081 GACAAGCCGGGGTTATAACCTTACATCAACCTTAGGCTCGAGCAAGGCCCGAAGGTGGTT 1140  
DB 1081 AGCAAGCCGGGGTTATAACCTTACATCAACCTTACGCTCAAGCAAGCCGGAAGGTGGTT 1140  
QY 1141 TATGTCGCAAGTTCCGGGCGGTCATCCAGGATATCTGTAACCTTCGAGGCTGG 1200  
DB 1141 TATGTCGCAAGTTCCGGGCGGTCATCCAGGATATCTGTAACCTTCGAGGCTGG 1200  
QY 1201 GTAGACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
DB 1201 GTAGACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGCACTATGGCTTACCTCCGTTGGAGCTACTCGGGGTGGGCTGA 1311  
DB 1261 GCCAAGCGGCACTATGGCTTACCTCCGTTGGAGCTATTGGGCTGA 1311

Search completed: June 29, 2004, 09:14:42  
Job time : 578 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 09:04:55 ; Search time 3765 Seconds  
(without alignments)  
10398.225 Million cell updates/sec

Title: US-10-081-739A-1

Perfect score: 1311

Sequence: 1 atggccaaactactccagact.....gctactgcgggtgggctga 1311

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107.6	8.2	540	9	AUI92670
2	82.8	6.3	428	13	BU639413
3	76.8	5.9	1628	28	BZ424843
4	72.8	5.6	394	13	BU640392

5	71.2	5.4	768	28	AQ159619
6	63.2	4.8	660	14	CD900121
7	62.6	4.8	681	29	CNS0100L
8	62.4	4.8	858	29	CG302842
9	60.8	4.6	691	29	CG213284
10	58	4.4	519	14	CB884205
11	56.6	4.3	422	9	AL820356
12	54	4.1	593	29	CG289556
13	54	4.1	602	29	CG345046
14	54	4.1	608	29	CG222587
15	54	4.1	695	28	CC344571
16	53.6	4.1	559	14	CF322783
17	53	4.0	405	14	CF323742
18	52.8	4.0	562	12	BJ477953
19	52.6	4.0	558	14	CF323939
20	52.4	4.0	474	10	AW923088
21	52	4.0	743	28	AQ159694
22	51.6	3.9	841	12	BI949299
23	51.2	3.9	613	29	CC639582
24	51.2	3.9	982	28	CC424389
25	48.8	3.7	681	29	CNS02EOD
26	46.6	3.6	424	10	BF177257
27	46.6	3.6	451	12	BG556788
28	46.6	3.6	532	12	BG557843
29	46.6	3.6	541	10	BF177143
30	46.6	3.6	542	10	BE361447
31	46.6	3.6	545	12	BG355757
32	46.6	3.6	615	10	BE360090
33	46.6	3.6	936	29	CG038303
34	46.2	3.5	500	9	AA898808
35	45.8	3.5	474	10	BE249153
36	45.8	3.5	584	14	CF554700
37	45.8	3.5	869	14	CK159167
38	45.6	3.5	740	13	BQ801516
39	44.8	3.4	525	14	CA700958
40	44.8	3.4	527	10	BE293754
41	44.8	3.4	834	29	CG133420
42	44.8	3.4	1201	13	BX356664
43	44.6	3.4	404	10	AW923211
44	44.6	3.4	642	14	CF306714
45	44.6	3.4	1201	13	BX381961

ALIGNMENTS

RESULT 1  
AUI92670  
LOCUS AUI92670 540 bp mRNA linear EST 14-OCT-2003  
DEFINITION Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis  
CDNA clone PFL034d01\_r 5', mRNA sequence.  
ACCESSION AUI92670  
VERSION AUI92670.1 GI:31931544  
KEYWORDS EST.  
SOURCE Porphyra yezoensis  
ORGANISM Porphyra yezoensis  
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.

REFERENCE 1 (bases 1 to 540)  
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.  
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS  
JOURNAL J. Phycol. 39 (5), 923-930 (2003)  
COMMENT Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/. Location/Qualifiers  
1. 540  
/organism="Porphyra yezoensis"

```

/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL034d01.r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

```

## ORIGIN

```

Query Match      8.2%; Score 107.6; DB 9; Length 540;
Best Local Similarity 58.4%; Pred. No. 1.1e-13;
Matches 188; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 30 CGGGGTCATAATGCAGCCCTTCTACTGGACGTGCTTCAGGAGAAATGTTGGACAC 89
Db 217 CGGGGTCCTCTCCCACTACTTCTACTGGACCTTCCCGCGGATGGAGCCTGTGGAAGAA 276

QY 90 AATACGGCAGAGATACCGGAGTGTGTACGATCCGGAATCTCCGCAATATGATTCCTCCC 149
Db 277 CTTGGCGGTGAACGGGTGGACCTCGCGCGCGGGGCTCACTGCGTCTGGCTGGCC 336

QY 150 GCGGACAGGGCATGGCGGGCGCTTTCATGATGGCTACGACCCCTAGACTTCTTTGA 209
Db 337 GCGGTACAAGGGCGAGCGCGGGCGCGCGGACGTGGCTATGGCTGTACGACACGTATGA 396

QY 210 CTTGGGTGATGACGACCAAGAGGAGCGGTAGAGACGGCTTTGGCTCCAGCAGGAGCT 269
Db 397 CTTGGGGAGTTTGACCAAAAGGGCACCGTACGGACCAAGTACGACCAAGGACGCT 456

QY 270 CGTGAACATGATAAACACCGCCACCGCTATGCGATGAAGTAAATAGCGGATATAGTCAT 329
Db 457 CTTGTCCGCGCGCTCGCGCGCGCATGACGCGGCGCATCCAGGTGTACGCGGCAATTGTCT 516

QY 330 CAACACCGCGCGCGGGGTGAC 351
Db 517 CAACACCGGGTGGCGCGAGAC 538

```

```

RESULT 2
BU639413
LOCUS      428 bp mRNA linear EST 06-MAY-2003
DEFINITION mgcwl3xN24f.b RCW Lambda Zap Express Library Magnaporthe grisea
ACCESSION  BU639413
VERSION     BU639413.2 GI:30396874
KEYWORDS   EST.
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS    Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
           1. (bases 1 to 428)
           Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P., Lu,G.,
           Bhatterai,K. and Dean,R.A.
TITLE      Expressed sequence tags from the rice blast fungus, Magnaporthe
           grisea
JOURNAL    Unpublished (2002)
COMMENT    On Sep 30, 2002 this sequence version replaced gi.23351739.
           Contact: Ebbole D
           Department of Plant Pathology & Microbiology
           Texas A&M University
           Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
           Tel: 979 845 4831
           Fax: 979 845 6483
           Email: d-ebbole@tamu.edu
           Chromatogram file of this sequence is available, see contact
           person:Best nr hit (April. 22, 2003) gb|EAA33974.1| hypothetical
           protein [Neurospora crassa] 178 1e-44
PCR PRIMERs
FORWARD: T3 primer
BACKWARD: T7 primer
Seq primer: T3.
           location/Qualifiers
           1. 428

```

```

/organism="Magnaporthe grisea"
/mol_type="mRNA"
/strain="Cp987"
/db_xref="taxon:149305"
/clone="mgcwl3xN24"
/sex="Mau1-1 hermaphrodite"
/tissue_type="Mycelium"
/dev_stage="Day 5 post-inoculation"
/clone_lib="RCW Lambda Zap Express Library"
/notes="Vector: pBluescript excised from Lambda Zap
Express; Site 1: EcoRI; Site 2: XhoI; Day 5
post-inoculation MNAs prepared from Magnaporthe grisea
grown at 23C in the dark with constant gyratory shaking
100 rpm in Vogel's minimal medium containing 0.5% isolated
rice cell walls as the sole carbon source. Library
provided by Sheng-Cheng Wu. Sequences were processed by
one of two methods. Where a full-length alignment to the
M. grisea genome sequence was available, the EST sequence
was trimmed according to the alignment, otherwise sequence
quality was assessed using phredPhrap version 991019 and
trimmed according to phd files (0.05) and for vector
seqs."

```

## ORIGIN

```

Query Match      6.3%; Score 82.8; DB 13; Length 428;
Best Local Similarity 55.5%; Pred. No. 4.2e-08;
Matches 181; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 37 ATATGCGAGCGTTCTACTGGACGTGCTTCAGGAGGAATATGTGGACACAAATACGG 96
Db 63 ATATGCGAGCGTTCTACTGGATGTATGCCAGCAGATCAGAAACACTGGTCTAGCTTGA 122

QY 97 CAGAATACATCCGAGTGTACGATCCGGAATCTCCGCAATATGGATTCCCGCGGAGC 156
Db 123 AAGAGATCCCCAGCTCAATCATGGGTATGCAATATCTGGTTCCACAGGATGC 182

QY 157 AAGGGATGGGGCGGCGCTTATTCGATGGGTACGACCCCTACGACTTTTTCACCTCGT 216
Db 183 AAGGGCTCGTCCAAGACCGCGC---AATGGGTACGACATCTATGACCTCTATGACCTGGGC 239

QY 217 GAGTACGACCAAGAGGACCGGTAGAGACCGCTTGGCTCCAGCAGCAGCTCGTGAAC 276
Db 240 GAGTTCGACCAAAAGGAAAGCGTGGCCACCAATGGGAAACCAAGGAGAGCTCGTCAAG 299

QY 277 ATGATAAACACCCGCCACGCTATGCGATGAAGGTAAATAGCCGATATAGTCATCAACAC 336
Db 300 CTTTGTCTACGGCCAAAGCCAAAGCGTTGGCATCTACTGGGACCGGTTCTGAACAC 359

QY 337 CGCGCGGGGTGACCTGGAGTGAA 362
Db 360 AGGTTTTCAGCGCACCAATAAGGAGAA 385

```

## RESULT 3

```

BU624843/c
LOCUS      1628 bp DNA linear GSS 13-DEC-2002
DEFINITION 100020232-5732 Aspergillus terreus random genomic DNA clone library
Aspergillus terreus genomic, genomic survey sequence.
ACCESSION  BU624843
VERSION     BU624843.1 GI:26666298
KEYWORDS   GSS.
SOURCE     Aspergillus terreus
ORGANISM   Aspergillus terreus
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
AUTHORS    Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
           1. (bases 1 to 1628)
           Askenazi,M., Driggers,E.M., Holtzman,D.A., Norman,T.C., Iverson,S.,
           Zimmer,D.P., Boers,M-E., Blomquist,P.R., Martinez,E.J.,
           Monreal,A.W., Feibelman,T.P., Mayorga,M.E., Maxon,M.E., Sykes,K.,
           Tobin,J., Cordero,E., Salama,S.R., Trueheart,J., Royer,J.C. and
           Madden,K.T.
           Integrating transcriptional and metabolite profiles to direct the
           engineering of lovastatin-producing strains

```

## FEATURES

```

source
1. 428

```

Peterson Bldg, MS2132, College Station, TX 77843-2132, USA  
Tel: 979 845 4831  
Fax: 979 845 6483  
Email: d-ebbole@tamu.edu  
Chromatogram file of this sequence is available, see contact  
person:best nr hit (April. 22. 2003) gb|EAA33974.1| hypothetical  
protein [Neurospora crassa] 155 9e-38  
PCR Primers  
FORWARD: T7 primer  
BACKWARD: T7 primer  
Plate: mgcw019 row: O column: 09  
Seq primer: T3

FEATURES  
source  
1. .394  
/organism="Magnaporthe grisea"  
/mol\_type="mRNA"  
/strain="CP987"  
/db\_xref="taxon:148305"  
/clone="mgcw019x009"  
/sex="Mat1-1 hermaphrodite"  
/tissue\_type="Mycelium"  
/dev\_stage="Day 5 post-inoculation"  
/clone\_lib="RCW Lambda Zap Express Library"  
/note="Vector: pBluescript excised from Lambda Zap  
Express; Site 1: EcoRI; Site 2: XhoI; Day 5  
post-inoculation mRNAs prepared from Magnaporthe grisea  
grown at 23C in the dark with constant gyratory shaking  
100 rpm in Vogel's minimal medium containing 0.5% isolated  
rice cell walls as the sole carbon source. Library  
provided by Sheng-Cheng Wu. Sequences were processed by  
one of two methods. Where a full-length alignment to the  
M. grisea genome sequence was available, the EST sequence  
was trimmed according to the alignment, otherwise sequence  
quality was assessed using phredphrap version 991019 and  
trimmed according to phd files (0.05) and for vector  
seqs."

ORIGIN  
Query Match 5.6%; Score 72.8; DB 13; Length 394;  
Best Local Similarity 56.9%; Pred. No. 7.4e-06;  
Matches 153; Conservative 0; Mismatches 113; Indels 3; Gaps 1;  
QY 37 ATAATCAGCGGCTTCTACTGGGACGTCCTTCAGGAGGAATATGTTGGGACACAAATACGG 96  
DB 91 ATGATCAGGGCTTTTGGTGGTATGTCTCCAGCAGATCAGAAACACTGGGTTCAGGCTTGA 150  
QY 97 CAGAGATACCGGAGTGTACGATGCCGAATCTCCGAATATGGATTCCTCCCGCGGAGC 156  
DB 151 AAGGAGATCCCCAGCTCANATCATGGGTATCGACAAATATCTGGGTTCACAGGATGC 210  
QY 157 AAGGCGATGGGGCGGCTTATTCGATGGGCTACGACCCCTACGACTTCTTTGACCTCGGT 216  
DB 211 AAGGGCTCTTCCAAGACCGGC---AATGGGTACGACATCTATGACCTCCATGCGGC 267  
QY 217 GAGTACGACGAGGAGGACGCTAGAGACGCGCTTTGGCTTCAAGCAGGAGCTCGTGAAC 276  
DB 268 GAGTTCGACCAAAAGGAGGAGCGTGGCCACCAATGTTGTAACCAAGAGGAGCTCGTCAAG 327  
QY 277 ATGATAAACACCGCCCGCCACGCTTATGGCAT 305  
DB 328 CTTTGTCTTACGGTCAAGGCCAAAGCGGT 356

RESULT 5  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AQ159619 768 bp DNA linear GSS 09-SEP-1998  
mgxb0001K09r CUGI Rice Blast BAC Library Magnaporthe grisea genomic  
clone mgxb0001K09r, genomic survey sequence.  
AQ159619 GI:3556644  
KEYWORDS  
Magnaporthe grisea (anamorph: Pyricularia grisea)  
Magnaporthe grisea

Unpublished (2002)  
Contact: Zimmer DP  
Microbia, Inc.  
One Kendall Square Building 1400 W, Cambridge, MA 02139, USA  
Tel: 617-621-8322  
Fax: 617-  
Email: dzimmer@microbia.com  
Class: plasmid ends.  
Location/Qualifiers  
1. .1628  
/organism="Aspergillus terreus"  
/mol\_type="genomic DNA"  
/strain="ATCC 20542 (A. terreus Thom, anamorph)"  
/db\_xref="taxon:33178"  
/lab\_host="Escherichia coli"  
/clone\_lib="Aspergillus terreus random genomic DNA clone  
library"  
/note="Vector: pZEROTM-2; Site 1: Sau3A; Site 2: BamHI;  
Sau3A genomic fragments ligated into BamHI digested  
pZEROTM-2"

ORIGIN  
Query Match 5.9%; Score 76.8; DB 28; Length 1628;  
Best Local Similarity 54.7%; Pred. No. 1.7e-06;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 33 GGTATATGACGGCTTCTACTCGGACGTCCTTCAGGAGGAATATGTTGGGACACAAT 92  
DB 317 GCTAATCTTCCAGGGTTTGAATGGCAGCTCCCGGACGCGGAGCACTGGCGGCGCT 258  
QY 93 ACGCAGAGATACCGGAGTGGTACGATCGCGGAATCTCCGAATATGGATTCCCGCGGC 152  
DB 257 CGCGCGGCATTCGCGGATTTGAGGCCATCGGAATGCACTACATATGGCTTCCCGCGG 198  
QY 153 GAGCAAGGCGATGGCGGCGCTTATCGATGGGTACGACCCCTACGACTTCTTTGACCT 212  
DB 197 ATGCAAGGGATGGACC---CATCGGGGACCGGTATGATATCTATGATGACTT 141  
QY 213 CGGTGAGTACACAGCAGAGGACCGGTAGACAGCGCTTTGGCTCCAGCAGGAGCTCGT 272  
DB 140 GGGGAGTTCATCAGAAAGCTTATCAGCAGCAAAATGGGGGACACGGGAGGAGCTGCA 81  
QY 273 GAAATGATAAACACCGCCCGCCAGCTATGGCATGAAGGTATATGCGGATAGTCATCAA 332  
DB 80 CGAGCTAGTTAACGACGCTCAGCCCTCGGCTGGGGTCTATTGGGACCGGTTCTCAA 21  
QY 333 CCACGCGCGCGGTGACC 352  
DB 20 CCACAAGGCTGTGCGGATC 1

RESULT 4  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BU640392 394 bp mRNA linear EST 06-MAY-2003  
mgcw019x009f.b RCW Lambda Zap Express Library Magnaporthe grisea  
cDNA clone mgcw019x009 5', mRNA sequence.  
BU640392  
BU640392.2 GI:30395965  
EST.  
Magnaporthe grisea (anamorph: Pyricularia grisea)  
Magnaporthe grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.  
1 (bases 1 to 394)  
Ebbole D.J., Yuan J., Thomas T.L., Bobrowicz P., Lu G.,  
Bhatterai K. and Dean R.A.  
Expressed sequence tags from the rice blast fungus, Magnaporthe  
grisea  
Unpublished (2002)  
On Sep 30, 2002 this sequence version replaced gi:23352718.  
Contact: Ebbole DJ  
Department of Plant Pathology & Microbiology  
Texas A&M University



```

ORGANISM  Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE 1 (bases 1 to 681)
AUTHORS   Direct Submission
TITLE     Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
JOURNAL   BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
REFERENCE 2 (bases 1 to 681)
AUTHORS   Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
TITLE     Direct Submission
JOURNAL   Submitted (16-FEB-2000) BSMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
COMMENT   This clone is from an A. Gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.
FEATURES             Location/Qualifiers
     source          1..681
                     /organism="Anopheles gambiae"
                     /mol_type="genomic DNA"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="26N13"
                     /clone_lib="NotreDame1"
                     /note="end : T7"
ORIGIN
Query Match      4.8%; Score 62.6; DB 29; Length 681;
Best Local Similarity 54.8%; Pred. No. 0.0019;
Matches 166; Conservative 0; Mismatches 134; Indels 3; Gaps 2;
QY 36 CATAAATCAGCGCTTACTCGGACGTCCTTACGAGGAGATATGGTGGACACATACG 95
Db 302 CTTCTCGAATTTTCACATGTTACTACCCGACGGCGGAATACTGGCAGAGGCCGC 243
QY 96 GCAGAGATACCGAGTGTACGATGCGGAATCTCCGCAATATGGATTCCCGCGCGAG 155
Db 242 CGAGCGCGCCCGCACCTGGCGGAATGGGCAATACCGACTGT-GCTCGGCTGCCTA 184
QY 156 CAAGGCATGGCGCGCTTTCGATGGGTAGACCCCTACGACTTCTTTGACCTCGG 215
Db 183 CAAGGGCGCTTCCGGCGGCTACTCCGTCGGCTAGATACTTACGATCTGTTCGATCTCGG 124
QY 216 TGAGTACACACAGAGGAGCGGTAGACGCGCTTTGGGTCCAAAGCAGGAGTCTGTA 275
Db 123 CGAATTTGACCAAGAGCGCGTGGCCACCAATACGGGCAAGCGCGCTCGAACA 64
QY 276 --CATGATAACCGCCCGCTATGGCATGAGGTAATAGCCGATATAGTCATCAAC 333
Db 63 CGCTCGGCAACTTCTCGCGCGAACACGGGTGGGACACTCTACGACGTGGTGTTCAC 4
QY 334 CAC 336
Db 3 CAC 1
RESULT 8
CG302842
LOCUS      CG302842            858 bp      DNA      linear      GSS 25-AUG-2003
DEFINITION CG302842TV_ZM_0.7_1.5_KB_Zea mays genomic clone ZMMBma0742C19,
genomic survey sequence.
ACCESSION  CG302842
VERSION    CG302842.1  GI:34217056
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 858)

```

```

AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE     Consortium for Maize Genomics
JOURNAL   Unpublished (2002)
COMMENT   Other GSSs: OG2A322TH
Contact: Cathy Whitelaw
TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..858
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBma0742C19"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match      4.8%; Score 62.4; DB 29; Length 858;
Best Local Similarity 47.4%; Pred. No. 0.0023;
Matches 186; Conservative 0; Mismatches 206; Indels 0; Gaps 0;
QY 253 GGTCTCAAGCAGGAGCTCGTGAACATGATAACACCGCCACGCCCTATGGCATGAGGTA 312
Db 223 GGCACAGGAGCGCAGCTCAAGTCCCTGATCGAGGGGTTCCACAGCAAGGCGTGAAGTTC 282
QY 313 ATAGCGGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGAACCCCTTCGTG 372
Db 283 ATCGCGGACATGTCATCAACCCACCGCGCGGAGCACCAGCGCGCGGCTCTTAC 342
QY 373 AAGGACTATACCTGGACCGACTTCTCAAAGTTCGGGTGCGGTAAATACACGCCCAACTAC 432
Db 343 TGCATCTTCAGGGCGGCGACCGCGGACTCCCGCTCGACTGGGGCCCCACATGATCTGC 402
QY 433 CTCGACTTCCACCGAAGAGCTCCATCGGGCGGATTCGGAAACATTTGGAGGCTATCCC 492
Db 403 AGCGACGACAGGCGCTACTCCGACGGCACGGGGAACCCCGACACCGCGCGCGCTTCGGC 462
QY 493 GACATATGCCAGACAGAGCTGGGACCGAGTACTGGCTCTGGCCAGCAGGAGGCTAC 552
Db 463 GCGGCGCGGACATCGACCTGAACCCGCGCTCCAGCAGGAGCTCACCGGCTGGCTC 522
QY 553 GCGGCATATCTCAGGAGCATCGCATCGCTCGCGCTTCGACTAGCTCAAGGGCTAT 612
Db 523 AACTGGCTCAAGACGACGTCGGCTTCGACGCGTGGCGCTCGACTTCGCCAAGGCTAC 582
QY 613 GTCCTCGGTCTGTCAGGACTGGCTGAAC TG 644
Db 583 TCCGCCGACGTCGCCAAGGTTCTACATCGACGG 614
RESULT 9
CG213284
LOCUS      CG213284            691 bp      DNA      linear      GSS 22-AUG-2003
DEFINITION CG213284TV_ZM_0.7_1.5_KB_Zea mays genomic clone ZMMBma0702N11,
genomic survey sequence.
ACCESSION  CG213284
VERSION    CG213284.1  GI:34113172
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 691)

```

**AUTHORS** Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
**TITLE** Consortium for Maize Genomics  
**JOURNAL** Unpublished (2002)  
**COMMENT** Other\_GSSs: OG0EW78TV  
 Contact: Cathy Whitelaw  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

**FEATURES** Location/Qualifiers  
 1..691  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0702N11"  
 /clone\_lib="ZM 0.7 1.5 kb"  
 /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

**ORIGIN**  
 Query Match 4.6%; Score 60.8; DB 29; Length 691;  
 Best Local Similarity 47.2%; Pred. No. 0.0049;  
 Matches 185; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 253 GGCTCCACGAGGCTCGTGAACATGATAAACACCGCCCGCTATGGCATGAAGTA 312  
 DB 129 GGCACGAGCGCAGCTCAAGTCCCTGATCGAGGGGTTCCACAGCAAGGGGTGAAGTC 188  
 QY 313 ATAGCGGATATAGTANTCAACACCGCGCGCGTGTGACTGGAGTGAACCCCTTGTG 372  
 DB 189 ATCGCGGACATCGTCATCAACACCGCAGGAGGAGCACCAGCGCGCGGCATCTAC 248  
 QY 373 AACGACTATACCTGGACCGACTTCCTCAAAGTCCGGTGGGTAAATACACGCCCAACTAC 432  
 DB 249 TGAATCTTCAGGGCGGACGCGGACTCCCGCTCGACTGGGCGCCACATGATCTGC 308  
 QY 433 CTGACTTCCACCGCAACAGAGTCCATCGCGCGATTCGGGAACATTTGGAGGCTATCCC 492  
 DB 309 AGCGACGACGAGCGGTACTCCGACGCGCAGCGGGAACCCCGACACCGCGCGGCTTCGGC 368  
 QY 493 GACATATGCGACGACGAGCTGGGACAGTACTGGCTCTGGCGCAGCGAGAGCTAC 552  
 DB 369 GGGCGCGCGGATGATGACCACTGAACCGCGCGTCCAGAGAGCTCACCGGTGGCTC 428  
 QY 553 GCGGCATATCTCAGGAGCATCGGCATCGATGCTCGCGCTTCGACTAGCTCAAGGGCTAT 612  
 DB 429 AACTGSGCTCAAGACGAGCTGGCTTCGACGCGTGGCGCTTCGACTTCGCCAAGGCTAC 488  
 QY 613 GCTCCCTGGGTCTCAGGACTGGGTGACTG 644  
 DB 489 TCCGCGGACGCTCGCAAGGTCTACATCGACGG 520

**RESULT 10**  
**LOCUS** CB884205  
**DEFINITION** HV10M08; HV Hordeum vulgare subsp. vulgare cDNA clone HV10M08  
**ACCESSION** CB884205  
**VERSION** CB884205.1  
**KEYWORDS** EST.  
**SOURCE** Hordeum vulgare subsp. vulgare  
**ORGANISM** Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Hordeum.  
**REFERENCE** 1 (bases 1 to 519)

**AUTHORS**  
**TITLE**  
**JOURNAL**  
**COMMENT**

Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.  
 Barley ESTs from germinating seeds  
 Unpublished (2002)  
 Contact: Stein Nils  
 Molecular Markers Group, Department Genbank  
 Institute of Plant Genetics and Crop Plant Research (IPK)  
 Corrensstr. 3, 06466, Gatersleben, Germany  
 Tel: 039482-5522  
 Fax: 039482-5595  
 Email: stein@ipk-gatersleben.de  
 Insert Length: 519 Std Error: 0.00  
 Plate: 10 row: M column: 8  
 Seq primer: M13rev.

**FEATURES**  
 source

1..519  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="barke"  
 /sub\_species="vulgare"  
 /db\_xref="GABI:269134"  
 /db\_xref="taxon:112509"  
 /clone="HV10M08"  
 /tissue\_type="germinating seeds"  
 /dev\_stage="germinating seeds (48-96 h)"  
 /lab\_host="X110-Gold"  
 /clone\_lib="HV"  
 /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Roots were grown for two days on filter paper at room temperature. Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI).  
 NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable. Average insert size is 1 kb"

**ORIGIN**

Query Match 4.4%; Score 58; DB 14; Length 519;  
 Best Local Similarity 46.7%; Pred. No. 0.019;  
 Matches 184; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 253 GGCTCCACGAGGCTCGTGAACATGATAAACACCGCCCGCTATGGCATGAAGTA 312  
 DB 106 GGCACCAAGCGCAGCTCAAGTCCCTCATCGCGCGCTCCACGCGCAAGGCGTCAAGGCC 165  
 QY 313 ATAGCGGATATAGTATCAACACCGCGCGGTGACTGGAGTGAACCCCTTCGTG 372  
 DB 166 ATCGCGCATCTGATATCAACACCGCGCGGTGACTGGAGTGAACCCCTTCGTG 225  
 QY 373 AACGACTATACCTGGACCGACTTCCTCAAAGTGGCTCGGGTAAATACACGGCCAACTAC 432  
 DB 226 TGCATCTTCGAGGCGCGCACCCCGGACGCGCGCTCGACTGGGGCCCCACATGATCTGC 285  
 QY 433 CTGACTTCCACCGAAGAGCTCCATCGGGCGATTCCGGAACATTTGAGGCTATCCC 492  
 DB 286 CGCGACGACCGCGCTTACGCGGACGCGCACCGGCAACCCGACACCGCGCGGCTTCGGG 345  
 QY 493 GACATATGCCACGACAGAGCTGGGACCACTAGTACTGGCTCTGGGCCACGAGGAGCTAC 552  
 DB 346 GCGCGCGCGGACATCGACCACTCAACCCCGCGGTCCAGAGGAGCTCGTCGAGTGGCTC 405  
 QY 553 GCGGCATATCTAGGAGCATCGGCATCGATGCTGGCGCTTGGACTACGTCAGGGCTAT 612  
 DB 406 AACTGSGCTCAGGACCGACGCTCGGCTTCGACGCGTGGCGCTTCGACTTCGCCAAGGCTAC 465  
 QY 613 GCTCCCTGGGTCTCAAGGACTGGCTGAACCTGT 646  
 DB 466 TCCGCGGACGCTGCGCAAGATCTACGTCGACCGCT 499

**RESULT 11**

AL820356  
 LOCUS 422 bp mRNA linear EST 15-JUL-2002  
 DEFINITION AL820356 N:130 Triticum aestivum cDNA clone H07\_N130\_plate\_32, mRNA sequence.  
 ACCESSION AL820356  
 VERSION AL820356.1 GI:21831956  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.  
 REFERENCE 1 (bases 1 to 422)  
 AUTHORS Wilson, I., Beswick, R., Shepherd, S., Barker, G., Parker, J., Owen, P., Edwards, D., Coghill, J., Holdsworth, M., Lenton, J., Shewry, P. and Edwards, K.  
 TITLE A BSRF-funded wheat EST resource for the academic community  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Barker G  
 Institute of Arable Crop Research  
 Long Ashton, Bristol BS41 9AF United Kingdom.  
 FEATURES  
 Location/Qualifiers  
 1..422  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="mercia"  
 /db\_xref="taxon:4565"  
 /clone="H07\_N130\_plate\_32"  
 /tissue\_type="embryo"  
 /dev\_stage="2 days post germination"  
 /clone\_lib="N:130"

ORIGIN  
 Query Match 4.3%; Score 56.6; DB 9; Length 422;  
 Best Local Similarity 47.4%; Pred. No. 0.035;  
 Matches 170; Conservative 0; Mismatches 189; Indels 0; Gaps 0;  
 QY 253 GGCTCCAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTATGGCTGAAGGTA 312  
 Db 62 GGCACAAAGGCGAGCTCAAGTCCCTGATCGGGCGCTCCACGCAAGGCGGTCAAGCC 121  
 QY 313 ATAGCCATATAGTCATCAACCCAGCGCGCGGTGACCTGGAGTGGAAACCCCTTGGTG 372  
 Db 122 ATCCGCCGACATCGTCATCAACCCAGCGCGCGGTGACCTGGAGTGGAAACCCCTTGGTG 372  
 QY 373 AACGACTATACCTGGACGAGCTTCTCAAGGTGCGTGGTAAATACACGGCCAACTAC 432  
 Db 182 TGCATCTTCGAGGCGGACCCCTGACGCGGCTCGACTGGGCCCCACATGATCTGC 241  
 QY 433 CTCGACTTCACACGAGCTCCATCGCGGGGATTCGGAACATTTGAGGCTATCCC 492  
 Db 242 CGCGACGACCGGCCCTACCGCAGCGGCAACCCAGACACCGGCGCGGACTTCGGG 301  
 QY 493 GACATATGCCACGACAGAGCTGGGACGAGTCTGGCTCTGGGCCAGCAGGAGAGCTAC 552  
 Db 302 GCGCCCGGACATGACCCACTCAACCCGCGGTCCAGAGGAGTCTGCGAGTGGCTC 361  
 QY 553 CGCGCATATCTCAGGAGCATCGGATCGCTGCGGCTTCGACTACGTCAGGAGGCTA 611  
 Db 362 AACTGGCTCAGGACCGAGCTCGGCTTCGACGCTGGGCTTCGACTTCGCAAGGCTA 420

RESULT 12  
 CG289556/c  
 LOCUS  
 DEFINITION CG289556 593 bp DNA linear GSS 25-AUG-2003  
 OGSAR78TC ZM 0.7.1.5\_KB Zea mays genomic clone ZMMBMA0819M12,  
 genomic survey sequence.  
 ACCESSION CG289556  
 VERSION CG289556.1 GI:34203770  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 593)  
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
 TITLE Consortium for Maize Genomics  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Cathy Whitelaw  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TF  
 Class: sheared ends.  
 FEATURES  
 Location/Qualifiers  
 1..593  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0819M12"  
 /clone\_lib="ZM 0.7.1.5\_KB"  
 /note="Vector: pBCK-; Site 1: HindII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

ORIGIN  
 Query Match 4.1%; Score 54; DB 29; Length 593;  
 Best Local Similarity 48.2%; Pred. No. 0.16;  
 Matches 184; Conservative 0; Mismatches 195; Indels 3; Gaps 1;  
 QY 264 GGAGCTCGTGAACATGATAAACAACCGCCACGCTATGGCTGAAGTAAATAGCGATAT 323  
 Db 422 GGAGCTCAAGTCCCTGATCGCGCGTTCACGCAAGGCGGTGAGGTGGCGGACAT 363  
 QY 324 AGTCATCAACACCGCGCGGTGACCTGGAGTGGAAACCCCTTGTGAACGACTATAC 383  
 Db 362 GTCATCAACACCGGTGCGCGACTACAGGACCGCGCGGCACTACTGTGATCTTGA 303  
 QY 384 CTGGACCGACTTCTCAAGGTGCGTGGGTAAATACACGGCCAACTACTCGACTTCCA 443  
 Db 302 GGGCGGACCGCGCGGCTGGACTGGGCCCCACATGATCTCGCGGACGACAC 243  
 QY 444 CCGGACGAGCTCCATCGGGCGATTCGGACATTTGGAGGCTATCCGACATATGCA 503  
 Db 242 CATATCTCGGACGCGGCGCACTCGACACCGGCGCGGCTTCGCGCGGCGCGCA 183  
 QY 504 ---CGACAAGAGCTGGACCACTGCTGCTGGGCCAGCCAGGAGAGCTACGCGGCATA 560  
 Db 182 CATCGACCACTCAACGACCGGCTCCAGGGGAGCTCACCGACTGGCTCTGGCTCAA 123  
 QY 561 TCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTACGTCAGGAGGCTATGCTCCCTG 620  
 Db 122 GTCGTCGACCTGGCTTCGACGCTGGCGCTCGACTTCGCAAGGCTACTCCTCCGCA 63  
 QY 621 GGTCTCGAAGGACTGGCTGAAC 642  
 Db 62 GGTCCCAAGGTCTACGTGCAC 41

RESULT 13  
 CG345046  
 LOCUS  
 DEFINITION CG345046 602 bp DNA linear GSS 26-AUG-2003  
 OGVD111TH ZM 0.7.1.5\_KB Zea mays genomic clone ZMMBMA0509B21,  
 genomic survey sequence.  
 ACCESSION CG345046  
 VERSION CG345046.1 GI:34262312  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1. (bases 1 to 602)  
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
JOURNAL  
COMMENT

Other GSSs: OGVDELITV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

Location/Qualifiers

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1. (bases 1 to 608)  
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
JOURNAL  
COMMENT

Other GSSs: OGVF50TV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

Location/Qualifiers

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1. (bases 1 to 608)  
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
JOURNAL  
COMMENT

Other GSSs: OGVF50TV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

Location/Qualifiers

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0565I03"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .608

/organism="Zea mays"

/mol\_type="genomic DNA"

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1. (bases 1 to 602)  
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
JOURNAL  
COMMENT

Other GSSs: OGVDELITV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

Location/Qualifiers

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0509B21"

/clone\_lib="ZM 0.7 1.5 KB"

/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

FEATURES

source

1. .602

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0



clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 695)  
 White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 Contact: Cathy Whitelaw  
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: Tg  
 Class: sheared ends.

FEATURES  
 source  
 1..695  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0343J11"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## ORIGIN

Query Match 4.1%; Score 54; DB 28; Length 695;  
 Best Local Similarity 48.2%; Pred. No. 0.17;  
 Matches 184; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

Qy 264 GGAGCTCGTGACATGATAACACCGCCCGCTATGGCATGAAGTAATAGCCGATAT 323  
 Db 192 GGAGCTCAAGTCCCTGATCGCCGCTTCCACGGCAAGGGGTGCAGGTGGTGGGACAT 251

Qy 324 AGTCATCAACACCGCCGCGGTGACCTGGAGTGAACCCCTTCGTGAACGACTATAC 383  
 Db 252 CGTCATCAACACCGGTGGCGGCTACTACAGGAGCCCGGGCATCTACTGCATCTCGA 311

Qy 384 CTGACCGACTTCTCAAGTTCGGTGGGTAAATACACGGCCAACTACCTGACTTCCA 443  
 Db 312 GGGCGGACCGCCGACCGCGCTGGACTGGGGCCCCACATGATCTGCCGGCAGCACAC 371

Qy 444 CCCGAACGAGCTCCATCGGGCGATTCCGGAACATTTGGAGGCTATCCGACATATGCCA 503  
 Db 372 CATATATCGGACGGCAGGCCACCTCGACACCGGGCGGACTTCGGCGGGCGCCGA 431

Qy 504 ---CGACAAGAGCTGGGACCACTAGTGGCTCTGGGCCAGCCAGGAGCTACCGGGCATA 560  
 Db 432 CATCGACCACTCAACGACCGCTCCAGCGGAGCTCACCGACTGGCTGCTCTGGCTCAA 491

Qy 561 TCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTAGTCAAGGGCTATGCTCCCTG 620  
 Db 492 GTCGTCGACCTGGGCTTCGACGCTGGCGCTGGCTGACTTCGCCAAGGGCTACTCCCCCGA 551

Qy 621 GGTCGTCGAAGGACTGGCTGAAC 642  
 Db 552 GGTGCCCAAGGTCTACGTCGAC 573

Search completed: June 29, 2004, 11:46:31  
 Job time : 3772 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 09:04:58 ; Search time 113 Seconds  
(without alignments)  
6438.413 Million cell updates/sec

Title: US-10-081-739A-1

Perfect score: 1311  
Sequence: 1 atggccaagtactccgagct.....gctactgcgggtgggtctga 1311

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.6	6.1	1458	4	US-09-291-023A-3
2	79.6	6.1	1458	4	US-09-540-715A-3
3	76.8	5.9	1683	3	US-08-814-052-5
4	76.8	5.9	1683	3	US-08-812-829-5
5	76.8	5.9	1777	1	US-08-146-422-33
6	76.8	5.9	1777	1	US-08-233-575-1
7	76.8	5.9	1777	1	US-08-626-554-16
8	76.8	5.9	1968	1	US-08-468-700-33
9	76.8	5.9	1968	1	US-08-645-971-1
10	76.8	5.9	1968	1	US-08-468-220-31
11	76.8	5.9	1968	2	US-08-468-698-31
12	76.8	5.9	1968	2	US-08-704-706A-33
13	76.8	5.9	1968	3	US-08-890-383-1
14	76.8	5.9	1968	3	US-08-914-679A-1
15	76.8	5.9	1968	3	US-08-985-659-33
16	76.8	5.9	1968	3	US-08-194-664A-31
17	76.8	5.9	1968	5	PCT-US94-01553A-31
18	76.8	5.9	1968	5	PCT-US95-10426-31
19	76.8	5.9	2149	1	US-07-623-953-2
20	76.8	5.9	2253	3	US-08-814-052-1
21	76.8	5.9	2253	3	US-08-812-829-1
22	76.4	5.8	1458	4	US-09-291-023A-1
23	76.4	5.8	1458	4	US-09-540-715A-1
24	75.2	5.7	1893	3	US-08-814-052-7
25	75.2	5.7	1893	3	US-08-812-829-7
26	75.2	5.7	1918	3	US-09-264-037-1
27	75.2	5.7	1920	1	US-08-720-899-1

28	75.2	5.7	1920	1	US-08-459-610-1	Sequence 1, Appli
29	75.2	5.7	1920	2	US-08-343-804-1	Sequence 1, Appli
30	75.2	5.7	1920	2	US-08-687-399-1	Sequence 1, Appli
31	75.2	5.7	1920	2	US-08-600-908A-1	Sequence 1, Appli
32	75.2	5.7	1920	3	US-08-683-838A-1	Sequence 1, Appli
33	75.2	5.7	1920	3	US-09-182-859-1	Sequence 12, Appli
34	75.2	5.7	1920	3	US-09-170-670-12	Sequence 12, Appli
35	75.2	5.7	1920	3	US-09-193-068-12	Sequence 12, Appli
36	75.2	5.7	1920	4	US-09-290-734-12	Sequence 12, Appli
37	75.2	5.7	1920	4	US-09-537-168-3	Sequence 3, Appli
38	75.2	5.7	1920	4	US-09-672-459-1	Sequence 1, Appli
39	75.2	5.7	1920	4	US-09-636-252A-1	Sequence 1, Appli
40	75.2	5.7	1920	4	US-09-545-586-12	Sequence 12, Appli
41	75.2	5.7	1920	4	US-10-186-042-1	Sequence 1, Appli
42	75.2	5.7	1920	4	US-09-769-864-12	Sequence 12, Appli
43	75.2	5.7	1926	3	US-09-183-412-12	Sequence 12, Appli
44	74.6	5.7	1443	4	US-09-537-168-1	Sequence 1, Appli
45	73.6	5.6	5677	1	US-07-623-953-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-291-023A-3  
; Sequence 3, Application US/09291023A  
; Patent No. 6309871  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic Ac  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/291,023A  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: DK 1999 00438  
; PRIOR FILING DATE: 1999-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 1458  
; TYPE: DNA  
; ORGANISM: Bacillus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1458)  
US-09-291-023A-3

Query Match	6.1%;	Score 79.6;	DB 4;	Length 1458;
Best Local Similarity	54.9%;	Pred.No. 1.1e-12;		
Matches 179;	Conservative	0;	Mismatches 144;	Indels 3; Gaps 1;
QY	25	AAGCGCGGGGTCAATAATCAGGCGGTTCTACTGGGACGTGCCCTTCAGGAGGAATATGGTGG	84	
Db	13	ACGAACGAACGATTATGCACTATTTTGAATGGAACGTTCCGGAATGATGCAACACATTGG	72	
QY	85	GACACAAATACGCGCAGAAGATACCGGAGTGGTACGATGCGGGAATCTCCGCAATATGGATT	144	
Db	73	AACCGCTTACACAAACGCTCAAAATTTAAAAATGCGGAAATACAGCAATCTGGATT	132	
QY	145	CCCCCGCGGACAGGGCATGGCGGCGCTTATTCGATGGCTACGACCCCTACGACTTC	204	
Db	133	CCACTTCGCTGGAAGGAACGAGCCAAATGATG---TAGGCTACGCTGCGTATGACCTT	189	
QY	205	TTTGACCTCGGTGAGTACGACCAAGGGAACGGTAGACGCGCTTTGGCTTCAAGCAG	264	
Db	190	TATGACCTTGGTGAATTTTAAACCAAAAGGAACGGTCCGTACGAAATATGGACAAAGCA	249	
QY	265	GAGCTCGTGAACATGATATAAACCGCCCGCTATGGCATGAAAGGTAAATAGCCGATATA	324	

Db 250 GAATTAGAACGAGCGATTGTTTCGTTAAAGGCAACGGGATTCAAGTGTATGGCGATGTT 309

QY 325 GTCAATCAACACCGCGCGCGGTGA 350

Db 310 GTTATGAACCAATAAAGCGGAGCTGA 335

## RESULT 2

US-09-540-715A-3  
; Sequence 3, Application US/09540715A  
; Patent No. 6623948  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic A  
; TITLE OF INVENTION: Encoding Same  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/540,715A  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/291,023  
; PRIOR FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 3  
; LENGTH: 1458  
; TYPE: DNA  
; ORGANISM: Bacillus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1458)  
US-09-540-715A-3

Query Match 6.1%; Score 79.6; DB 4; Length 1458;  
Best Local Similarity 54.9%; Pred. No. 1.1e-12;  
Matches 179; Conservative 0; Mismatches 144; Indels 3; Gaps 1;  
QY 25 AAGCGCGGGTCAATGATGAGCGGTTCTACTGGACGTCCTTCAAGGAGGAATATGGTGG 84  
Db 13 ACCAAGCGGAACGATTATGAGTATTTTGAATGGAACGTTCCGGAATGATGGAACAATGG 72  
QY 85 GACACAATACGCGAGAGATACCGAGTGGTACGATCGCGAATCTCCGCAATATGGATT 144  
Db 73 AACCGTTACACAACACGCTCAAAATTTAAATGCGGAATTTACAGCAATCTGGATT 132  
QY 145 CCCCCGGGAGCAAGGCGATCGGCGCGCCTATTTCGATGGGCTACGACCCCTACGACTTC 204  
Db 133 CCACCTGCGTGAAGAGAACGAGCCAAATATG--TAGGCTACGCTGCGTATGACCTT 189  
QY 205 TTTCACCTCGGTAGTACGACGAGGAGGAGGAGGAGCGCTTTGGCTCCAAAGCAG 264  
Db 190 TATGACCTTGGTGAATTTAAACCAAAAGGAGCGGTCGCTACCAATATGGAACAAAGCA 249  
QY 265 GAGCTCGTGAACATGATAAACACCGCCCATGTCGATGAGGATGAAGGTAAATAGCCGATATA 324  
Db 250 GAATTAGACGAGCGATTTCGTTCTTAAAGGCAACGGGATTCAGTGTATGGCGATGTT 309  
QY 325 GTCAATCAACACCGCGCGCGGTGA 350  
Db 310 GTTATGAACCAATAAAGCGGAGCTGA 335

## RESULT 3

US-08-814-052-5  
; Sequence 5, Application US/08814052  
; Patent No. 6015783  
; GENERAL INFORMATION:  
; APPLICANT: von der Osten, Claus  
; APPLICANT: Cherty, Joel R.  
; APPLICANT: Bjornvad, Mads E.  
; APPLICANT: Vind, Jesper

; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING  
; TITLE OF INVENTION: OR STAINS FROM CELLULOSIC FABRIC  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6015783o No. 6015783disk of No. 6015783th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/814,052  
; FILING DATE: 06-MAR-1997  
; CLASSIFICATION: 510  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4684.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:

; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1683 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-814-052-5

Query Match 5.9%; Score 76.8; DB 3; Length 1683;  
Best Local Similarity 54.7%; Pred. No. 7.2e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCATATGACGCGTCTTACTGGACGTCCTTCCAGGAGGAATATGGTGGACACA 90  
Db 100 GGGACGCTGATGCGATTTTGAATGCTATGCTCCCAATGCGGCCCAACATTTGGAAGCGT 159  
QY 91 ATACGCGAGAGATACCGGAGTGGTACGATCGCGAATCTCCGCAATATGGATTTCCTCCG 150  
Db 160 TTGCAAAACGACTCGGCATATTTGGCTGAACAGGATTTACTGCGCTCTGGATTCCCTCG 219  
QY 151 GCAGCAAGGCGATGGCGCGCTTATTCGATGGGCTACGACCCCTACGACTTCTTTGAC 210  
Db 220 GCATATAAGGAGACGAGCAAGCGGATG--TGGCTACGGTGTCTTACGACCTTTATGAT 276  
QY 211 CTCGGTGTAGTACGACGAGGAGGAGGAGGAGCGCTTTGGCTCCAGGAGGAGCTC 270  
Db 277 TTAGGGAGGTTTTCATCAAAAGGAGCGGTTCCGACAAAGTACGGCAACAAAGAGAGCTG 336  
QY 271 GTGAACATGATAAACACCGCCCATGTCGATGAAGTAAATAGCCGATATAGTCATC 330  
Db 337 CAATCTGCGATCAAAAGTCTTCATTCCTCCGCGACATTAAGCTTTACGGGATGTGGTCATC 396  
QY 331 AACCAACCGCGCGCGGTGA 350  
Db 397 AACCAAAAGCGCGCTGA 416

## RESULT 4

US-08-812-829-5  
; Sequence 5, Application US/08812829  
; Patent No. 6017751  
; GENERAL INFORMATION:  
; APPLICANT: von der Osten, Claus  
; APPLICANT: Bjornvad, Mads E.

APPLICANT: Wind, Jesper  
APPLICANT: Rasmussen, Michael Dolberg  
TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING  
TITLE OF INVENTION: OR STAINS FROM CELLULOSIC FABRIC  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 60177510 No. 6017751disk of No. 6017751th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/812,829  
FILING DATE: 06-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4690.204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1683 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-812-829-5

Query Match 5.9%; Score 76.8; DB 3; Length 1683;  
Best Local Similarity 54.7%; Pred. No. 7.2e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATGACGAGGCTTCTACTGGGAGCTGCTTTCAGGAGGAATATGTTGGACACA 90  
Db 100 GGGAGCTGATGCAATATTTTGAATGTACATGCCCAATGACGCGCAACATTGGAAGCGT 159  
QY 91 ATACGGCAGAGATACCGGAGTGTAGTCCGGAATCTCCGCAATATGATTCCTTTGAC 150  
Db 160 TTGCAAAACGACTCGGCATATTTGGCTGAACACCGGTATTACTGCGCTCGGATTCCTCCCG 219  
QY 151 GCGAGCAGGCGATGGGCGGCGCTATTTCGATGGGCTACGACCGCTTCTTTGAC 210  
Db 220 GCATATAGGGAACGAGCCAGCGATG---TGGGCTACGGTGTACGACCTTTATGAT 276  
QY 211 CTCGGTGTAGTACGACCAAGAGGGAACCGGTAGAGACGCGCTTTGGCTCCAGCAGGAGCTC 270  
Db 277 TTAGGGAGGTTTCATCAAAAAGGAGCGGTTCCGCAAAAGTACGCAAAAAGGAGAGCTG 336  
QY 271 GTGACATGATAACACCGCCCGCTATGCGATGAGGTATAGCCGATATAGTCATC 330  
Db 337 CAATCTCGGATCAAAAGTCTTCATTCGCGGACATTAACGTTTACGGGATGTGTCATC 396  
QY 331 AACACCGCGCGCGGCTGA 350  
Db 397 AACCAAAAGGCGGCTGA 416

RESULT 5  
US-08-146-422-33  
; Sequence 33, Application US/08146422  
; Patent No. 5543576  
; GENERAL INFORMATION:  
; APPLICANT: VAN OOLJEN, ALBERT J. J.

APPLICANT: RIETVELD, KRIJN  
APPLICANT: HOEKEMA, ANDREAS  
APPLICANT: PEN, JAN  
APPLICANT: SIJMONS, PETER C.  
APPLICANT: VERWOERD, TEUNIS C.  
APPLICANT: QUAX, WILHEMUS J.  
TITLE OF INVENTION: PRODUCTION OF ENZYMES IN SEEDS AND THEIR  
TITLE OF INVENTION: USE  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/146,422  
FILING DATE: 02-NOV-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: KENNEDY, BILL  
REGISTRATION NUMBER: 33,407  
REFERENCE/DOCKET NUMBER: 44615-20011.23  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-146-422-33

Query Match 5.9%; Score 76.8; DB 1; Length 1777;  
Best Local Similarity 54.7%; Pred. No. 7.4e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATGACGAGGCTTCTACTGGGAGCGTCTTCAGGAGGAATATGTTGGACACA 90  
Db 109 GGGAGCTGATGCAATATTTTGAATGTGTACATGCCCAATGACGCGCAACATTGGAAGCGT 168  
QY 91 ATACGGCAGAGATACCGGAGTGTAGTCCGGAATCTCCGCAATATGATTCCTTTGAC 150  
Db 169 TTGCAAAACGACTCGGCATATTTGGCTGAACACCGGTATTACTGCGCTCTGGATTCCTCCCG 228  
QY 151 GCGAGCAGGCGATGGGCGGCGCTATTTCGATGGGCTAGGACCGCTTCTTTGAC 210  
Db 229 GCATATAGGGAACGAGCGCGATG---TGGGCTACGGTGTACGACCTTTATGAT 285  
QY 211 CTCGGTGTAGTACGACCAAGAGGGAACCGGTAGAGACGCGCTTTGGCTCCAGCAGGAGCTC 270  
Db 286 TTAGGGAGTTCATCAAAAAGGAGCGGTTCCGCAAAAGTACGCGCAAAAGGAGAGCTG 345  
QY 271 GTGACATGATAACACCGCCCGCTATGCGATGAGGTATAGCCGATATAGTCATC 330  
Db 346 CAATCTCGGATCAAAAAGTCTTCATTCGCGGACATTAACGTTTACGGGATGTGTCATC 405  
QY 331 AACACCGCGCGCGGCTGA 350  
Db 406 AACCAAAAGGCGGCTGA 425

RESULT 6  
US-08-253-575-1  
; Sequence 1, Application US/08253575

Patent No. 5705375  
GENERAL INFORMATION:  
APPLICANT: VAN COYEN, ALBERT J.J.  
APPLICANT: RIETVELD, KRIJN  
APPLICANT: QUAX, WILHELMUS J.  
APPLICANT: VAN DEN ELZEN, PETRUS J.M.  
APPLICANT: PEN, JAN  
APPLICANT: HOEKEMA, ANDREAS  
APPLICANT: SIMONS, PETER C.  
TITLE OF INVENTION: TRANSGENIC PLANTS HAVING A MODIFIED  
CARBOHYDRATE CONTENT  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/253,575  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/849,422  
FILING DATE: 12-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20033.00  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-253-575-1

Query Match 5.9%; Score 76.8; DB 1; Length 1777;  
Best Local Similarity 54.7%; Pred. No. 7.4e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCATATGAGCGGTTCTACTGGGACGTCCTTCAGGAGGAATATGGTGGGACACA 90  
DB 109 GGGACGCTGATGCGATGTTTGAATGTTACATGCGCAATGACGGCCAACTTGGAAAGGT 168  
QY 91 ATACGGCAGAGATACCGGAGTGGTACGATGCGGAACTCCGCAATATGATTCCTCCCG 150  
DB 169 TTGCAAAACGATCGCGCATATTTGGCTGAACACGGTATTACTGCGCTCGGATTCCTCCCG 228  
QY 151 GCGAGCAGGCGATGGCGCGCTTATTGATGGGTACGACCCCTACGACTTCTTTGAC 210  
DB 229 GCATATAGGGAACGAGCCAAAGCGGATG---TGGGCTACGGTCTTACGACCTTATGAT 285  
QY 211 CTCGGTGTAGTACGACGAGGGAACGGTAGAGACGCGCTTTGGCTTCCAGCAGAGCTC 270  
DB 286 TTAGGGGAGTTTTCATCAAAAGGAGCGGTTCCGCAAAAAGTAGTCGGCACAAGGAGAGCTG 345  
QY 271 GTGAACATGATAAACCACCGCCACCGCTATGCGATGAAGGTAATAGCGATATAGTCATC 330  
DB 346 CAATCTGCGATCAAAAGTCTTCTCCCGGACATTAACGTTTACGGGATGTGTCATC 405  
QY 331 AACCAACCGCGCGCGGTGA 350  
|||||

DB 406 AACCAAAAGGGCGCTGA 425  
RESULT 7  
US-08-626-554-16  
Sequence 16, Application US/08626554  
Patent No. 5714474  
GENERAL INFORMATION:  
APPLICANT: VAN COIJEN, ALBERT J.J.  
APPLICANT: RIETVELD, KRIJN  
APPLICANT: HOEKEMA, ANDREAS  
APPLICANT: PEN, JAN  
APPLICANT: SIMONS, PETER C.  
APPLICANT: VERMOREL, TEUNIS C.  
TITLE OF INVENTION: PRODUCTION OF ENZYMES IN SEEDS AND THEIR  
TITLE OF INVENTION: USE  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 PENNSYLVANIA AVENUE NW  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/626,554  
FILING DATE: 02-APR-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 26192-20011.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030 MRSNFORWSH  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-626-554-16

Query Match 5.9%; Score 76.8; DB 1; Length 1777;  
Best Local Similarity 54.7%; Pred. No. 7.4e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCATATGAGCGGTTCTACTGGGACGTCCTTCAGGAGGAATATGGTGGGACACA 90  
DB 109 GGGACGCTGATGCGATGTTTGAATGTTACATGCGCAATGACGGCCAACTTGGAAAGGT 168  
QY 91 ATACGGCAGAGATACCGGAGTGGTACGATGCGGAACTTCGCAATATGATTCCTCCCG 150  
DB 169 TTGCAAAACGATCGCGCATATTTGGCTGAACACGGTATTACTGCGCTCGGATTCCTCCCG 228  
QY 151 GCGAGCAGGCGATGGCGCGCTTATTGATGGGTACGACCCCTACGACTTCTTTGAC 210  
DB 229 GCATATAGGGAACGAGCCAAAGCGGATG---TGGGCTACGGTCTTACGACCTTATGAT 285  
QY 211 CTCGGTGTAGTACGACGAGGGAACGGTAGAGACGCGCTTTGGCTTCCAGCAGAGCTC 270  
DB 286 TTAGGGGAGTTTTCATCAAAAGGAGCGGTTCCGCAAAAAGTAGTCGGCACAAGGAGAGCTG 345  
QY 271 GTGAACATGATAAACCACCGCCACCGCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
DB 346 CAATCTGCGATCAAAAGTCTTCTCCCGGACATTAACGTTTACGGGATGTGTCATC 405

QY 331 AACACCGCGCGCGGTGA 350  
|||||  
Db 406 AACCAAAAGCGCGCTGA 425

## RESULT 8

US-08-468-700-33  
; Sequence 33, Application US/08468700  
; Patent No. 5736499  
; GENERAL INFORMATION:

; APPLICANT: COLIN MITCHINSON  
; APPLICANT: CAROL A. REQUADT  
; APPLICANT: TRACI H. ROPP  
; APPLICANT: LEIF P. SOLHEIM  
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE  
; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genencor International

; STREET: 180 Kimball Way

; CITY: South San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468,700

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 252

; ATTORNEY/AGENT INFORMATION:

; NAME: Stone, Christopher

; REGISTRATION NUMBER: 35,696

; REFERENCE/DOCKET NUMBER: GC275

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 742-7555

; TELEFAX: (415) 742-7217

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1968 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-468-700-33

Query Match 5.9%; Score 76.8; DB 1; Length 1968;

Best Local Similarity 54.7%; Pred. No. 7.7e-12;

Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 31 GGGGTCTAATGCGCGCTTCTACTGGACGCTGCTTCAGGAGGAATATGTTGGACACA 90  
|||  
Db 263 GGGACGCTGATGCAGTATTTGAATGTTACATGCTCCCAATGACGCGCAACATTGGAAGCGT 322

QY 91 ATACGGCAGACATACCGAGTGTACGATCCGGAATCTCGCAATATGATTTCCCGG 150  
|||  
Db 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTAAGTCTGCGCTGCGGCTGAT 382

QY 151 GCGAGCAAGGCGATGGCGGCGCTTATTCGATGGCTACGACCCCTACGACTTCTTTGAC 210  
|||  
Db 383 GCATATAAGGAACAGCAGCAAGCGGATG--TGGCTACGGTGTCTAGGACTTTATGAT 439

QY 211 CTCGCTGATGACGACGAGGAAACGGTAGAGACGCGCTTTGGCTCCAGCAGAGCTC 270  
|||  
Db 440 TTAGGGAGTTTTCATCAAAAAGGGACGGTTTCGCAAAAGTAGTCGGCACAAGAGAGAGTG 499

QY 271 GTGAACATGATAAACACCGCCCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
|||  
Db 500 CAATCTCGGATCAAAAGTCTTCATTTCCCGGACATTAACGTTTACGGGATGTGTCATC 559

QY 331 AACACCGCGCGCGGTGA 350  
|||||  
Db 560 AACCAAAAGCGCGCTGA 579

## RESULT 9

US-08-645-971-1

; Sequence 1, Application US/08645971

; Patent No. 5763385

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Modified Alpha-Amylases Having Altered

; TITLE OF INVENTION: Calcium Binding Properties

; NUMBER OF SEQUENCES: 5

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/645,971

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1968 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Bacillus licheniformis

US-08-645-971-1

Query Match 5.9%; Score 76.8; DB 1; Length 1968;

Best Local Similarity 54.7%; Pred. No. 7.7e-12;

Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 31 GGGGTCTAATGCGCGCTTCTACTGGACGCTGCTTCAGGAGGAATATGTTGGACACA 90  
|||  
Db 263 GGGACGCTGATGCAGTATTTGAATGTTACATGCTCCCAATGACGCGCAACATTGGAAGCGT 322

QY 91 ATACGGCAGACATACCGAGTGTACGATCCGGAATCTCGCAATATGATTTCCCGG 150  
|||  
Db 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTAAGTCTGCGCTGCGGCTGAT 382

QY 151 GCGAGCAAGGCGATGGCGGCGCTTATTCGATGGCTACGACCCCTACGACTTCTTTGAC 210  
|||  
Db 383 GCATATAAGGGAACGAGCAAGCGGATG--TGGGCTACGCTGCTTACGACCTTTATGAT 439

QY 211 CTCGCTGATGACACCAAGGGAACGGTAGAGACGCGCTTTGGCTCCAGCAGGAGCTC 270  
|||  
Db 440 TTAGGGAGTTTTCATCAAAAAGGGACGGTTTCGCAAAAGTAGTCGGCACAAGAGAGAGCTG 499

QY 271 GTGAACATGATAAACACCGCCCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
|||  
Db 500 CAATCTCGGATCAAAAGTCTTCATTTCCCGGACATTAACGTTTACGGGATGTGTCATC 559

QY 331 AACACCGCGCGCGGTGA 350  
|||||  
Db 560 AACCAAAAGCGCGCTGA 579

## RESULT 10

US-08-468-220-31

; Sequence 31, Application US/08468220

; Patent No. 5824532

; GENERAL INFORMATION:

; APPLICANT: Antrim, Richard L.

; APPLICANT: Barnett, Christopher

; APPLICANT: Mitchinson, Colin

; APPLICANT: Power, Scott D.

; APPLICANT: Reguadt, Carol

; APPLICANT: Solheim, Leif P.

; TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase

NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,220  
FILING DATE: 06-JUN-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/194,664  
FILING DATE: 10-FEB-94  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,395  
FILING DATE: 11-FEB-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Stone, Christopher L.  
REGISTRATION NUMBER: 35,696  
REFERENCE/DOCKET NUMBER: GC220D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7555  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1968 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-468-220-31

Query Match 5.9%; Score 76.8; DB 1; Length 1968;  
Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATATGACGGCTTCTACTGGAGCGTCTTACGAGGAGTAATGTTGGGACACA 90  
DB 263 GGGACGCTGATGACGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 322  
QY 91 ATACGGCAGAGATACCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 150  
DB 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCGCTCTGATTCGCCCG 382  
QY 151 GCGAGCAGGCTATGGCGGCGCTATTTCGATGGGTACGACCCCTACGACTTCTTTGAC 210  
DB 383 GCATATAGGAGACGACGACGAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439  
QY 211 CTCGGTCACTACGACGAGGAAACGGTAGACGCGCTTTGGCTTCAAGCAGGAGCTC 270  
DB 440 TTAGGGGAGTTTCATCAAAAAGGAGCGGTTCCGACAAAGTACGACCAAAAGGAGAGCTG 499  
QY 271 GTGAACATGATTAACACCGCCCGCTATGCGATGAGGTATAGCCGATATAGTCATC 330  
DB 500 CAATCTCGCATCAAAAGTCTTCAATCCCGCATTAACGTTTACGGGATGTGTCATC 559  
QY 331 AACACCGCGCGCGCGTGA 350  
DB 560 AACCAAAAGGCGCGCTGA 579

RESULT 11  
US-08-468-698-31  
Sequence 31, Application US/08468698  
Patent No. 5845549  
GENERAL INFORMATION:

APPLICANT: Antrim, Richard L.  
APPLICANT: Barnett, Christopher  
APPLICANT: Mitchinson, Colin  
APPLICANT: Power, Scott D.  
APPLICANT: Requa, Carol  
APPLICANT: Solheim, Leif P.  
TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,698  
FILING DATE: 06-JUN-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/194,664  
FILING DATE: 10-FEB-94  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,395  
FILING DATE: 11-FEB-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Stone, Christopher L.  
REGISTRATION NUMBER: 35,696  
REFERENCE/DOCKET NUMBER: GC220D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7555  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1968 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-468-698-31

Query Match 5.9%; Score 76.8; DB 2; Length 1968;  
Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATATGACGGCTTCTACTGGAGCGTCTTACGAGGAGTAATGTTGGGACACA 90  
DB 263 GGGACGCTGATGACGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 322  
QY 91 ATACGGCAGAGATACCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 150  
DB 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCGCTCTGATTCGCCCG 382  
QY 151 GCGAGCAGGCTATGGCGGCGCTATTTCGATGGGTACGACCCCTACGACTTCTTTGAC 210  
DB 383 GCATATAGGAGACGACGACGAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439  
QY 211 CTCGGTCACTACGACGAGGAAACGGTAGACGCGCTTTGGCTTCAAGCAGGAGCTC 270  
DB 440 TTAGGGGAGTTTCATCAAAAAGGAGCGGTTCCGACAAAGTACGACCAAAAGGAGAGCTG 499  
QY 271 GTGAACATGATTAACACCGCCCGCTATGCGATGAGGTATAGCCGATATAGTCATC 330  
DB 500 CAATCTCGCATCAAAAGTCTTCAATCCCGCATTAACGTTTACGGGATGTGTCATC 559  
QY 331 AACACCGCGCGCGCGTGA 350  
DB 560 AACCAAAAGGCGCGCTGA 579



RESULT 12  
US-08-704-706A-33  
; Sequence 33, Application US/08704706A  
; Patent No. 5958739  
; GENERAL INFORMATION:  
; APPLICANT: GENENCOR INTERNATIONAL, INC.  
; APPLICANT: COLIN MITCHINSON  
; APPLICANT: ANTHONY G. DAY  
; APPLICANT: ANDREW SHAW  
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International, Inc.  
; STREET: 925 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1013  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,706A  
; FILING DATE: February 20, 1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stone, Christopher  
; REGISTRATION NUMBER: 35,696  
; REFERENCE/DOCKET NUMBER: GC275-3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 846-7555  
; TELEFAX: (650) 845-6405  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1968 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-704-706A-33

Query Match 5.9%; Score 76.8; DB 2; Length 1968;  
Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
  
QY 31 GGGGTCAATAATGACGGCGTTCTACTGGGACGTGCTTACGAGGAAATATGTTGGGACACA 90  
DB 263 GGGACGCTGATGCATATTTTGAATGGTACATGCCCATGACGCCCAACATTGGAAGCGT 322  
  
QY 91 ATACGGCAGAGATACCGGAGTGTGTACGATGCCGGAATCTCCGCAATATGATTCCCGCG 150  
DB 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCCGCTTGGATTCCCGCG 382  
  
QY 151 GCGACCAAGGCGATGGCGGCGCTTATTCGATGGGTACGACCCCTTGGCTCCAGCAGGAGCTC 210  
DB 383 GCATATAAGGGAACGAGCCACGCGATG---TGGGCTACGGTGTCTTACGACCTTTATGAT 439  
  
QY 211 CTCGGTGTAGTACGACAGAGGAAACGGTAGAGACGCGCTTTGGCTCCAGCAGGAGCTC 270  
DB 440 TTAGGGGAGTTTCATCAAAAAGGGACGGTTCGGAACAAAGTACGGCACAAGAGGAGGTG 499  
  
QY 271 GTGAACATGATAAACACCGCCACGCTATGGCATGAAGGTAAATAGCCGATATAGTCATC 330  
DB 500 CAATCTGCGATCAAAAGTCTTCAITCCCGGACATTAAACGTTTACGGGGATGTGGTCATC 559  
331 AACCAACCGCGCGCGGTGA 350  
560 AACCAACAAAGGCGCGCTGA 579

RESULT 14  
US-08-914-679A-1  
; Sequence 1, Application US/08914679A

RESULT 13  
US-08-890-383-1  
; Sequence 1, Application US/08890383  
; Patent No. 6008026  
; GENERAL INFORMATION:  
; APPLICANT: Anthony G. Day  
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE HAVING INTRODUCED  
; TITLE OF INVENTION: THEREIN A DISULFIDE BOND  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International, Inc.  
; STREET: 925 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1013  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/890,383  
; FILING DATE: To Be Assigned  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stone, Christopher  
; REGISTRATION NUMBER: 35,696  
; REFERENCE/DOCKET NUMBER: GC377  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 846-7555  
; TELEFAX: (650) 845-6504  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1968 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-890-383-1

Query Match 5.9%; Score 76.8; DB 3; Length 1968;  
Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
  
QY 31 GGGGTCAATAATGACGGCGTTCTACTGGGACGTGCTTACGAGGAAATATGTTGGGACACA 90  
DB 263 GGGACGCTGATGCATATTTTGAATGGTACATGCCCATGACGCCCAACATTGGAAGCGT 322  
  
QY 91 ATACGGCAGAGATACCGGAGTGTGTACGATGCCGGAATCTCCGCAATATGATTCCCGCG 150  
DB 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCCGCTTGGATTCCCGCG 382  
  
QY 151 GCGACCAAGGCGATGGCGGCGCTTATTCGATGGGTACGACCCCTTGGCTCCAGCAGGAGCTC 210  
DB 383 GCATATAAGGGAACGAGCCACGCGATG---TGGGCTACGGTGTCTTACGACCTTTATGAT 439  
  
QY 211 CTCGGTGTAGTACGACAGAGGAAACGGTAGAGACGCGCTTTGGCTCCAGCAGGAGCTC 270  
DB 440 TTAGGGGAGTTTCATCAAAAAGGGACGGTTCGGAACAAAGTACGGCACAAGAGGAGGTG 499  
  
QY 271 GTGAACATGATAAACACCGCCACGCTATGGCATGAAGGTAAATAGCCGATATAGTCATC 330  
DB 500 CAATCTGCGATCAAAAGTCTTCAITCCCGGACATTAAACGTTTACGGGGATGTGGTCATC 559  
331 AACCAACCGCGCGCGGTGA 350  
560 AACCAACAAAGGCGCGCTGA 579

RESULT 14  
US-08-914-679A-1  
; Sequence 1, Application US/08914679A

Patent No. 6080568  
 GENERAL INFORMATION:  
 APPLICANT: Anthony G. Day  
 APPLICANT: Barbara A. Swanson  
 TITLE OF INVENTION: MUTANT ALPHA-AMYLASE COMPRISING MODIFICATION  
 TITLE OF INVENTION: AT RESIDUES CORRESPONDING TO A210, H405 AND/OR  
 TITLE OF INVENTION: T412 IN BACILLUS LICHENIFORMIS  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genencor International, Inc.  
 STREET: 925 Page Mill Road  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304-1013  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/914,679A  
 FILING DATE: To Be Assigned  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stone, Christopher  
 REGISTRATION NUMBER: 35,696  
 REFERENCE/DOCKET NUMBER: GC387  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 846-7555  
 TELEFAX: (650) 845-6504  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-914-679A-1

Query Match 5.9%; Score 76.8; DB 3; Length 1968;  
 Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
 Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
 QY 31 GGGGTCAATATGACGGGTTCTACTGGGACGTGCTTCTAGGAGGAATATGGTGGACACA 90  
 Db 263 GGGACGCTGATGTCAGTATTTTGAATGGTACATGCCCAATGACGGCCAAATGGAGCGT 322  
 QY 91 ATACGGCAGAAATACCGGAGTGTACGATGCGGAATCTCCGCAATATGGATTCCCGCG 150  
 Db 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTCCGCTCTGGATTCCCGCG 382  
 QY 151 GCGAGCAAGGGCATGGCGCGGCTTATTCGATGGGTACGACCCCTACGACTTCTTTGAC 210  
 Db 383 GCATATAAGGGAACGAGCCAAAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439  
 QY 211 CTCGGTGTAGTACGACGAGGGAACGGTAGACGCGCTTTGGCTCCAAAGCAGGAGCTC 270  
 Db 440 TTAGGGAGTTTTCATCAAAAGGACGCTTCGGACAAAGTACGGCACAAAGGAGAGCTG 499  
 QY 271 GTGAACATGATAAACACCGCCCGCTATGGCATGAAGGTAATAGCCGATATAGTCATC 330  
 Db 500 CAATCTCGCATCAAAAGTCTTCATTCCCGCGACATTAAACGTTTACGGGATGTGGTCATC 559  
 QY 331 AACACCGCGCGCGGTGA 350  
 Db 560 AACCAAAAGGCGCGCTGA 579

RESULT 15

US-08-985-659-33  
 ; Sequence 33, Application US/08985659  
 ; Patent No. 6211134  
 ; GENERAL INFORMATION:

APPLICANT: Caldwell, Robert M  
 APPLICANT: Mitchinson, Colin  
 APPLICANT: Ropp, Traci M  
 TITLE OF INVENTION: Mutant Alpha-Amylase  
 FILE REFERENCE: A-69396/DJB/DAV/JJD  
 CURRENT APPLICATION NUMBER: US/08/985,659  
 CURRENT FILING DATE: 1997-12-09  
 PRIOR APPLICATION NUMBER: 08/645,971  
 PRIOR FILING DATE: 1996-05-14  
 NUMBER OF SEQ ID NOS: 41  
 SOFTWARE: Patent in Ver. 2.1  
 SEQ ID NO 33  
 LENGTH: 1968  
 TYPE: DNA  
 ORGANISM: Bacillus licheniformis  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (164)..(1699)  
 US-08-985-659-33

Query Match 5.9%; Score 76.8; DB 3; Length 1968;  
 Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
 Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
 QY 31 GGGGTCAATATGACGGGTTCTACTGGGACGTGCTTCTAGGAGGAATATGGTGGACACA 90  
 Db 263 GGGACGCTGATGTCAGTATTTTGAATGGTACATGCCCAATGACGGCCAAATGGAGCGT 322  
 QY 91 ATACGGCAGAAATACCGGAGTGTACGATGCGGAATCTCCGCAATATGGATTCCCGCG 150  
 Db 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTCCGCTCTGGATTCCCGCG 382  
 QY 151 GCGAGCAAGGGCATGGCGCGGCTTATTCGATGGGTACGACCCCTACGACTTCTTTGAC 210  
 Db 383 GCATATAAGGGAACGAGCCAAAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439  
 QY 211 CTCGGTGTAGTACGACGAGGGAACGGTAGACGCGCTTTGGCTCCAAAGCAGGAGCTC 270  
 Db 440 TTAGGGAGTTTTCATCAAAAGGACGCTTCGGACAAAGTACGGCACAAAGGAGAGCTG 499  
 QY 271 GTGAACATGATAAACACCGCCCGCTATGGCATGAAGGTAATAGCCGATATAGTCATC 330  
 Db 500 CAATCTCGCATCAAAAGTCTTCATTCCCGCGACATTAAACGTTTACGGGATGTGGTCATC 559  
 QY 331 AACACCGCGCGCGGTGA 350  
 Db 560 AACCAAAAGGCGCGCTGA 579

Search completed: June 29, 2004, 11:48:28  
 Job time : 116 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 10:43:46 ; Search time 607 Seconds  
(without alignments)  
9929.476 Million cell updates/sec

Title: US-10-081-739A-1

Perfect score: 1311

Sequence: 1 atggccaagtactccagact.....gctactgcgggtgggtgtga 1311

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3031105 seqs, 2298700234 residues

Total number of hits satisfying chosen parameters: 6062210

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US04\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1311	100.0	1311	15	US-10-146-662-1
2	1311	100.0	1311	15	US-10-105-733-1
3	1311	100.0	1311	15	US-10-081-739A-1
4	1239	94.5	1311	15	US-10-081-872-5
5	1239	94.5	1311	16	US-10-385-305-5
6	1231	93.9	1311	15	US-10-081-872-71
7	1231	93.9	1311	16	US-10-385-305-71
8	1192.6	91.0	1311	15	US-10-081-872-37
9	1192.6	91.0	1311	16	US-10-385-305-37
10	1186.2	90.5	1311	15	US-10-081-872-39
11	1186.2	90.5	1311	16	US-10-385-305-39
12	1183	90.2	1311	15	US-10-081-872-29
13	1183	90.2	1311	16	US-10-385-305-29
14	1171.8	89.4	1311	15	US-10-081-872-13

15	1171.8	89.4	1311	16	US-10-385-305-13	Sequence 13, Appl
16	1160.6	88.5	1311	15	US-10-081-872-61	Sequence 61, Appl
17	1160.6	88.5	1311	16	US-10-385-305-61	Sequence 61, Appl
18	1155.8	88.2	1311	15	US-10-081-872-55	Sequence 55, Appl
19	1155.8	88.2	1311	16	US-10-385-305-55	Sequence 55, Appl
20	1147.8	87.6	1311	15	US-10-081-872-3	Sequence 3, Appl
21	1147.8	87.6	1311	16	US-10-081-872-25	Sequence 25, Appl
22	1147.8	87.6	1311	15	US-10-081-872-45	Sequence 45, Appl
23	1147.8	87.6	1311	16	US-10-385-305-3	Sequence 3, Appl
24	1147.8	87.6	1311	16	US-10-385-305-25	Sequence 25, Appl
25	1147.8	87.6	1311	16	US-10-385-305-45	Sequence 45, Appl
26	1146.2	87.4	1311	15	US-10-081-872-59	Sequence 59, Appl
27	1146.2	87.4	1311	16	US-10-385-305-59	Sequence 59, Appl
28	1144.6	87.3	1311	15	US-10-081-872-51	Sequence 51, Appl
29	1144.6	87.3	1311	16	US-10-385-305-51	Sequence 51, Appl
30	1143	87.2	1311	15	US-10-081-872-63	Sequence 63, Appl
31	1143	87.2	1311	16	US-10-385-305-63	Sequence 63, Appl
32	1141.4	87.1	1311	15	US-10-081-872-19	Sequence 19, Appl
33	1141.4	87.1	1311	16	US-10-385-305-19	Sequence 19, Appl
34	1139.8	86.9	1311	15	US-10-081-872-15	Sequence 15, Appl
35	1139.8	86.9	1311	16	US-10-385-305-15	Sequence 15, Appl
36	1138.2	86.8	1311	15	US-10-081-872-11	Sequence 11, Appl
37	1138.2	86.8	1311	16	US-10-385-305-11	Sequence 11, Appl
38	1130.2	86.2	1311	15	US-10-081-872-41	Sequence 41, Appl
39	1130.2	86.2	1311	16	US-10-385-305-41	Sequence 41, Appl
40	1120.6	85.5	1311	15	US-10-081-872-1	Sequence 1, Appl
41	1120.6	85.5	1311	16	US-10-385-305-1	Sequence 1, Appl
42	1119	85.4	1311	15	US-10-081-872-35	Sequence 35, Appl
43	1119	85.4	1311	16	US-10-385-305-35	Sequence 35, Appl
44	1117.4	85.2	1311	15	US-10-081-872-23	Sequence 23, Appl
45	1117.4	85.2	1311	15	US-10-081-872-31	Sequence 31, Appl

#### ALIGNMENTS

RESULT 1  
US-10-146-662-1  
; Sequence 1, Application US/10146662  
; Publication No. US2003003172A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerendash, Joel  
; TITLE OF INVENTION: NOVEL METHODS OF ENZYME PURIFICATION  
; FILE REFERENCE: 09010-109001  
; CURRENT APPLICATION NUMBER: US/10/146,662  
; CURRENT FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1308)  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-146-662-1

Query Match	100.0%	Score 1311	DB 15	Length 1311
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1311	Conservative 0	Mismatches 0	Indels 0	Caps 0
QY	1	ATGCCCAAGTACTCCGAGCTGGAAAGGGCGGGTCTAATGCGAGCGTTCTACTCGGAC	60	
Db	1	ATGCCCAAGTACTCCGAGCTGGAAAGGGCGGGTCTAATGCGAGCGTTCTACTCGGAC	60	
QY	61	GTCCCTTCAGGAGGAATATGGTGGGACACAATACGCCAGAAATACCGAGTGGTACGAT	120	
Db	61	GTCCCTTCAGGAGGAATATGGTGGGACACAATACGCCAGAAATACCGAGTGGTACGAT	120	
QY	121	GCCGGATCTCCGCAATATGGATTCCCCCGGCGGACGAGGCGCATGGCGCGCCTATTG	180	

Db 121 GCGGAAATCTCGCAATATGGATTCCTCCCGGGAGCAAGGGCATATGGCGCGCCCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTCTTTTGGACCTCGGTGAGTACGACCAAGGAAGCAAGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTCTTTTGGACCTCGGTGAGTACGACCAAGGAAGCAAGGTA 240  
Qy 241 GAGACGGCTTTGGCTCAGAGAGAGCTCGTGAACATATAAACACCGCCACGCTAT 300  
Db 241 GAGACGGCTTTGGCTCAGAGAGAGCTCGTGAACATATAAACACCGCCACGCTAT 300  
Qy 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTTCGGTTCGGGTAAATAC 420  
Db 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTTCGGTTCGGGTAAATAC 420  
Qy 421 ACGGCAACTACCTGACTTCCACCGAAGAGCTCCATGCGGGCGATTCCGGAACATTT 480  
Db 421 ACGGCAACTACCTGACTTCCACCGAAGAGCTCCATGCGGGCGATTCCGGAACATTT 480  
Qy 481 GGAGGCTATCCCGACATATGCGACGAGCTGGACCACTGCTTGGGCGAGC 540  
Db 481 GGAGGCTATCCCGACATATGCGACGAGCTGGACCACTGCTTGGGCGAGC 540  
Qy 541 CAGGAGCTACCGGCACTATCTCAGGAGCTGGCATCGATGCTGCGGCGCTTCGACTAC 600  
Db 541 CAGGAGCTACCGGCACTATCTCAGGAGCTGGCATCGATGCTGCGGCGCTTCGACTAC 600  
Qy 601 GTCAAGGCTATGCTCCTCGGTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
Db 601 GTCAAGGCTATGCTCCTCGGTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
Qy 661 GTTGGAGACTTGGGACACCAACGCTGAGCGCTGTTCTCAACTGGGCATATCTGAGCGGT 720  
Db 661 GTTGGAGACTTGGGACACCAACGCTGAGCGCTGTTCTCAACTGGGCATATCTGAGCGGT 720  
Qy 721 GCCAAGTCTTGGACTTGGCCCTCTACTCAAGATGGATGAGGCTTTCACCAACAAAC 780  
Db 721 GCCAAGTCTTGGACTTGGCCCTCTACTCAAGATGGATGAGGCTTTCACCAACAAAC 780  
Qy 781 ATTCAGCGCTGCTCTGCTGCTTCAAGCGGCGAGACTGTTCTCCGCGACCGCTTC 840  
Db 781 ATTCAGCGCTGCTCTGCTGCTTCAAGCGGCGAGACTGTTCTCCGCGACCGCTTC 840  
Qy 841 AAGCGCTAACCTTTGAGCAACCCAGCACCGATATATCTGGAACAGTATCCAGCC 900  
Db 841 AAGCGCTAACCTTTGAGCAACCCAGCACCGATATATCTGGAACAGTATCCAGCC 900  
Qy 901 TACCGCTTCACTCCTACCTACGAGGCGCGGACCAATATTTACCGCGACTACGAGGAG 960  
Db 901 TACCGCTTCACTCCTACCTACGAGGCGCGGACCAATATTTACCGCGACTACGAGGAG 960  
Qy 961 TGGCTCAACAGGTAAGCTCAAGAACTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGTAAGCTCAAGAACTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020  
Qy 1021 AGCACCGACATAGTCTACTACGATAAGATGAATCATCTTCTGTCAGGAACGGCTACGGG 1080  
Db 1021 AGCACCGACATAGTCTACTACGATAAGATGAATCATCTTCTGTCAGGAACGGCTACGGG 1080  
Qy 1081 GACAAAGCGGGGCTTAACTTACATCAACTAGGCTCGAGCAAGCGCGAAGGTGGTT 1140  
Db 1081 GACAAAGCGGGGCTTAACTTACATCAACTAGGCTCGAGCAAGCGCGAAGGTGGTT 1140  
Qy 1141 TATGTCGAGGTTGCGGGCGGTGATCCACGAGTATCTGTAACCTTCGAGGCTGG 1200  
Db 1141 TATGTCGAGGTTGCGGGCGGTGATCCACGAGTATCTGTAACCTTCGAGGCTGG 1200  
Qy 1201 GTAGACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCAGCTTACGACCT 1260

Db 1201 GTAGACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Qy 1261 GCCAACGGGAGTATGGCTACTCGGTGAGAGCTACTGCGGGGTGGCTGA 1311  
Db 1261 GCCAACGGGAGTATGGCTACTCGGTGAGAGCTACTGCGGGGTGGCTGA 1311  
RESULT 2  
US-10-105-733-1  
; Sequence 1, Application US/10105733  
; Publication No. US20030138786A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Miller, Carl  
; APPLICANT: Kazaoka, Martin  
; APPLICANT: Short, Jay  
; APPLICANT: Mathur, Eric  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 03010-107002  
; CURRENT APPLICATION NUMBER: US/10105,733  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 10/081,739  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated  
US-10-105-733-1  
Query Match 100.0%; Score 1311; DB 15; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCCAAGTACTCCGAGCTCGAAGGCGGGGTCTATAATGAGGCTTCTACTGGGAC 60  
Db 1 ATGCCAAGTACTCCGAGCTCGAAGGCGGGGTCTATAATGAGGCTTCTACTGGGAC 60  
Qy 61 GTGCTTCAGAGGATATGTTGGGACACATACGCGAGAGATACCGGAGTGTAGCAT 120  
Db 61 GTGCTTCAGAGGATATGTTGGGACACATACGCGAGAGATACCGGAGTGTAGCAT 120  
Qy 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCTATTTCG 180  
Db 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCTATTTCG 180  
Qy 181 ATGGCTACGACCCCTACGACTCTTTTGGACCTCGGTGAGTACGACCAAGGAAGCAAGGTA 240  
Db 181 ATGGCTACGACCCCTACGACTCTTTTGGACCTCGGTGAGTACGACCAAGGAAGCAAGGTA 240  
Qy 241 GAGACGGCTTTGGCTCAGAGAGAGCTCGTGAACATATAAACACCGCCACGCTAT 300  
Db 241 GAGACGGCTTTGGCTCAGAGAGAGCTCGTGAACATATAAACACCGCCACGCTAT 300  
Qy 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTTCGGTTCGGGTAAATAC 420  
Db 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTTCGGTTCGGGTAAATAC 420

```
QY 421 ACGGCACTACCTCGACTTCCACCGAAGAGCTCCATCGGGGATTCGGAAACATTT 480
Db 421 ACGGCACTACCTCGACTTCCACCGAAGAGCTCCATCGGGGATTCGGAAACATTT 480
QY 481 GGAGGCTATCCGACATATGCCACGACAGAGCTGGGACCACTGCTCTGGGCCAGC 540
Db 481 GGAGGCTATCCGACATATGCCACGACAGAGCTGGGACCACTGCTCTGGGCCAGC 540
QY 541 CAGGAGAGTACCGCGCATATCTCAGGAGATCGGCATCGATCGCTGGCGCTTGCATPAC 600
Db 541 CAGGAGAGTACCGCGCATATCTCAGGAGATCGGCATCGATCGCTGGCGCTTGCATPAC 600
QY 601 GTCAAGGGCTATGCTCCCTCGGCTCGTCAAGGACTCGCTGAAGTGTGGGAGGCTGGSCG 660
Db 601 GTCAAGGGCTATGCTCCCTCGGCTCGTCAAGGACTCGCTGAAGTGTGGGAGGCTGGSCG 660
QY 661 GTTGAGAGTACTGGGACCAACAGCTCGACGCTGTCTCAACTGGGCATACCTCGAGCGGT 720
Db 661 GTTGAGAGTACTGGGACCAACAGCTCGACGCTGTCTCAACTGGGCATACCTCGAGCGGT 720
QY 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGATGAGGCTTTGACAAACAAAC 780
Db 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGATGAGGCTTTGACAAACAAAC 780
QY 781 ATTCAGGCTCGTCTCTGCTGCTTCCAGAACGCGCAGACTGTTGCTCCCGCGACCCGTTTC 840
Db 781 ATTCAGGCTCGTCTCTGCTGCTTCCAGAACGCGCAGACTGTTGCTCCCGCGACCCGTTTC 840
QY 841 AAGGCGGTAACTTTGTAGCAACCAACGACACCGGATTAATCTGGAAACAGATTCAGGCC 900
Db 841 AAGGCGGTAACTTTGTAGCAACCAACGACACCGGATTAATCTGGAAACAGATTCAGGCC 900
QY 901 TACGCGTTTCATCTCCTCAGGAGGCGCAGCAATATTTCTACCGGACTACGAGGAG 960
Db 901 TACGCGTTTCATCTCCTCAGGAGGCGCAGCAATATTTCTACCGGACTACGAGGAG 960
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020
QY 1021 AGCAGGACATAGTCTACTACGATAAGATCAATCATCTCTGATGATGAGAACCTCGCGGAG 1080
Db 1021 AGCAGGACATAGTCTACTACGATAAGATCAATCATCTCTGATGATGAGAACCTCGCGGAG 1080
QY 1081 GACAAAGCGGGCTTTAATCACTACATCAACCTAGGCTCGAGCAAGGCGGAGGAGTGGGTT 1140
Db 1081 GACAAAGCGGGCTTTAATCACTACATCAACCTAGGCTCGAGCAAGGCGGAGGAGTGGGTT 1140
QY 1141 TATGTCGGAAGTTTCGGGGCGGTCATCCACGAGTATCTGTAACCTCGAGGCTGG 1200
Db 1141 TATGTCGGAAGTTTCGGGGCGGTCATCCACGAGTATCTGTAACCTCGAGGCTGG 1200
QY 1201 GTAGACAGTACGCTTACTCAAGGGGCTGGGCTCTATCTCGAAGCTCCAGCTTACGACCT 1260
Db 1201 GTAGACAGTACGCTTACTCAAGGGGCTGGGCTCTATCTCGAAGCTCCAGCTTACGACCT 1260
QY 1261 GCCAAGCGGAGTATGGCTACTCGTGTGGAGTACTCGGGGTGGGCTGA 1311
Db 1261 GCCAAGCGGAGTATGGCTACTCGTGTGGAGTACTCGGGGTGGGCTGA 1311
```

## RESULT 3

```
US-10-081-739A-1
; Sequence 1, Application US/1081739A
; Publication No. US20030170634A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Miller, Carl
; APPLICANT: Kazaoka, Martin
; APPLICANT: Short, Jay
```

```
; APPLICANT: Mathur, Eric
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; FILE REFERENCE: 09010-107001
; CURRENT APPLICATION NUMBER: US/10/081,739A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated
US-10-081-739A-1
```

```
Query Match 100.0%; Score 1311; DB 15; Length 1311;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCCCAAGTACTCCGAGCTGGAAAGGGGGGGTCAATATGAGCGCTTCTACTGGGAC 60
Db 1 ATGCCCAAGTACTCCGAGCTGGAAAGGGGGGGTCAATATGAGCGCTTCTACTGGGAC 60
QY 61 GTCCCTTCAGGAGGAATATGTTGGGACACAATACGGCAGAAATACCGAGTGGTACGAT 120
Db 61 GTCCCTTCAGGAGGAATATGTTGGGACACAATACGGCAGAAATACCGAGTGGTACGAT 120
QY 121 GCCGGAATCTCCGCAATATGGATTCCTCCGCGGAGCAAGGGCATGGGCGGCTATTCG 180
Db 121 GCCGGAATCTCCGCAATATGGATTCCTCCGCGGAGCAAGGGCATGGGCGGCTATTCG 180
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTTCGGTGTAGTACGACAGAGGGAACGGTA 240
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTTCGGTGTAGTACGACAGAGGGAACGGTA 240
QY 241 GAGACCGGCTTTGGTCTCAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTAT 300
Db 241 GAGACCGGCTTTGGTCTCAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTAT 300
QY 301 GGCAATGAAGTAAATACCGCATATAGTCATCAACCAACCGCGCGGCTGACCTGAGTGG 360
Db 301 GGCAATGAAGTAAATACCGCATATAGTCATCAACCAACCGCGCGGCTGACCTGAGTGG 360
QY 361 AACCCCTTCGTGAACGACTATACCTGGGACCGACTTCTCAAAGGTGCGCTCGGGTAAATAC 420
Db 361 AACCCCTTCGTGAACGACTATACCTGGGACCGACTTCTCAAAGGTGCGCTCGGGTAAATAC 420
QY 421 ACGGCACTACCTCGACTTCCACCGGAAGAGCTCCATGCGGGGATTCGGGAACATTT 480
Db 421 ACGGCACTACCTCGACTTCCACCGGAAGAGCTCCATGCGGGGATTCGGGAACATTT 480
QY 481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGGACCACTGCTGGCTCTGGGCCAGC 540
Db 481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGGACCACTGCTGGCTCTGGGCCAGC 540
QY 541 CAGGAGAGTACCGGCGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTGCATAC 600
Db 541 CAGGAGAGTACCGGCGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTGCATAC 600
QY 601 GTCAAGGGCTATGCTCCCTCGGCTCGTCAAGGACTCGCTGAAGTGTGGGAGGCTGGSCG 660
Db 601 GTCAAGGGCTATGCTCCCTCGGCTCGTCAAGGACTCGCTGAAGTGTGGGAGGCTGGSCG 660
QY 661 GTTGAGAGTACTGGGACCAACAGCTCGACGCTGTCTCAACTGGGCATACCTCGAGCGGT 720
Db 661 GTTGAGAGTACTGGGACCAACAGCTCGACGCTGTCTCAACTGGGCATACCTCGAGCGGT 720
```

OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-5

Query Match	Score 1239;	DB 15;	Length 1311;
Best Local Similarity	96.6%;	Pred. No. 0;	
Matches 1266;	Conservative 0;	Mismatches 45;	Indels 0;
Gaps 0;			
QY	1	ATGCCAAGTACTCCGAGCTGGAAAGGGGGGTCATAATCAGGCGTTCTACTGGGAC	60
DB	1	ATGCCAAGTACTCCGAGCTGGAAAGGGGGGTCATAATCAGGCGTTCTACTGGGAC	60
QY	61	GTGCTTTCAGAGAGATATGTTGGGACACAATACCGCAGAGATACCGAGTGGTACGAT	120
DB	61	GTCCAGGTGGAGGATCTGTTGGGACACCATCAGGAGCAAGATACCGGAGTGGTACGAG	120
QY	121	GCAGGAATCTCCGAGATATGGAATTCCTCCGCGCAGCAAGGCGATGGCGCGCTATTCG	180
DB	121	GCAGGAATCTCCGAGATATGGAATTCCTCCGCGCAGCAAGGCGATGGCGCGCTATTCG	180
QY	181	ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAAACGGTA	240
DB	181	ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAAACGGTA	240
QY	241	GAGACGCGTTTGGCTCCAAAGCAGGAGCTCGTGAAATGATGAAACACCGCCACGCTAT	300
DB	241	GAGACGCGTTTGGCTCCAAAGCAGGAGCTCGTGAAATGATGAAACACCGCCACGCTAT	300
QY	301	GGCATGAAGGTATAGCCGATATAGTCAATCAACCGCGCCCGCGGTGACCTGGAGTGG	360
DB	301	GGCATGAAGGTATAGCCGATATAGTCAATCAACCGCGCCCGCGGTGACCTGGAGTGG	360
QY	361	AACCCCTTCGTGAAACGACTATACCTGGACCGACTTCTTCAAGGCTCGCGTAAATAC	420
DB	361	AACCCCTTCGTGAAACGACTATACCTGGACCGACTTCTTCAAGGCTCGCGTAAATAC	420
QY	421	ACGCGCAACTACCTCGACTTCCACCGCAAGAGCTCCATCGCGCGGATTCGCGAACATTT	480
DB	421	ACGCGCAACTACCTCGACTTCCACCGCAAGAGCTCCATCGCGCGGATTCGCGAACATTT	480
QY	481	GGAGGCTATCCGACATATGCCACGACAGAGCTGGGACCAAGTACTGGCTTGGGCGAGC	540
DB	481	GGAGGCTATCCGACATATGCCACGACAGAGCTGGGACCAAGTACTGGCTTGGGCGAGC	540
QY	541	CAGGAGCTACGCGGATATCTCAGAGCATCGGATCGATCGCTGGCGCTTCGACTAC	600
DB	541	CAGGAGCTACGCGGATATCTCAGAGCATCGGATCGATCGCTGGCGCTTCGACTAC	600
QY	601	GTCAAGGCTATGCTCCCTGGTCTCAAGGCTCGGCTGAACTGGTGGGAGGCTGGGCG	660
DB	601	GTCAAGGCTATGCTCCCTGGTCTCAAGGCTCGGCTGAACTGGTGGGAGGCTGGGCG	660
QY	661	GTTGGAGAGTACTGGGACACCAACGCTCGACGCTGTTCTCAACTGGGATACTCGAGCGGT	720
DB	661	GTTGGAGAGTACTGGGACACCAACGCTCGACGCTGTTCTCAACTGGGATACTCGAGCGGT	720
QY	721	GCCAGGCTTGTGAGTTCGCTTACTACAGATGATGAGGCTTTCGACCAACAAAC	780
DB	721	GCCAGGCTTGTGAGTTCGCTTACTACAGATGATGAGGCTTTCGACCAACAAAC	780
QY	781	ATTCAGGCTGCTCTGCGCTTCAGAACCGCCAGACTGTTGTCTCCGCGACCCGCTTC	840
DB	781	ATTCAGGCTGCTCTGCGCTTCAGAACCGCCAGACTGTTGTCTCCGCGACCCGCTTC	840
QY	841	AAGCGGTAACTTGTGAGCAACCAAGACACCGATATATCTGGAACAAGTATCCAGCC	900
DB	841	AAGCGGTAACTTGTGAGCAACCAAGACACCGATATATCTGGAACAAGTATCCAGCC	900
QY	901	TACGGGTTTCACTCCCTACCTAGGCGGCGCAGCAATATTTCTACCGCGACTACGAGAG	960
DB	901	TATGCTTTTCACTCCCTACCTAGGCGGCGCAGCCCGTTCATATTTCTACCGCGACTACGAGAG	960
QY	961	TGGCTCAACAGGATAGCTCAAGACCTCATCTGGATACATGAACTCGCCGAGGA	1020
DB	961	TGGCTCAACAGGATAGCTCAAGACCTCATCTGGATACATGAACTCGCCGAGGA	1020
QY	1021	AGCACCGCATAGTCTACTACGATACGATGAACCTCATCTTCGTGAGAACCGCTACCGG	1080
DB	1021	AGCACCGCATAGTCTACTACGATACGATGAACCTCATCTTCGTGAGAACCGCTACCGG	1080
QY	1081	GACAACCGGGGTTTATACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	1140
DB	1081	GACAACCGGGGTTTATACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT	1140
QY	1141	TATGTCCGAAGTTCCGGCGCGGTGCATCCAGAGTACTGTTAACTCGGAGGCTGG	1200
DB	1141	TATGTCCGAAGTTCCGGCGCGGTGCATCCAGAGTACTGTTAACTCGGAGGCTGG	1200
QY	1201	GTAGAACGATGCTACTCAACGCGGTGCTTATCTCGAAGCTCCAGCTTACGACCT	1260
DB	1201	GTAGAACGATGCTACTCAACGCGGTGCTTATCTCGAAGCTCCAGCTTACGACCT	1260
QY	1261	GCCAAAGGCGAGTATGCTACTCCGCTGAGGCTACTCGGGGGTGGGCTGA	1311
DB	1261	GCCAAAGGCGAGTATGCTACTCCGCTGAGGCTACTCGGGGGTGGGCTGA	1311

RESULT 4  
US-10-081-872-5  
Sequence 5, Application US/10081872  
Publication No. US20030125534A1  
GENERAL INFORMATION:  
APPLICANT: Callen, Walter  
APPLICANT: Richardson, Toby  
APPLICANT: Frey, Gerhard  
APPLICANT: Short, Jay M.  
APPLICANT: Mathur, Eric J.  
APPLICANT: Gray, Kevin A.  
APPLICANT: Kerovuo, Janne S.  
TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
TITLE OF INVENTION: AND METHODS OF USE THEREOF  
FILE REFERENCE: 09010-108001  
CURRENT APPLICATION NUMBER: US/10/081.872  
CURRENT FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: US 60/270,495  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: US 60/270,496  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: US 60/291,122  
PRIOR FILING DATE: 2001-05-14  
NUMBER OF SEQ ID NOS: 321  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 1311  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:

961	TGGCTCAACAGGACAGGTTGAACAACTCATATGTATACACGACCACCTCGCAGGTGGA	1020
Db		
1021	AGCACCGACATAGTCTACTACGATAACGATGAATCATCTTCGTCAGGACGCTACGGG	1080
Qy		
1021	AGCACCGACATAGTCTACTACGATAACGATGAATCATCTTCGTCAGGACGCTACGGG	1080
Db		
1081	GACAAGCCGGGGCTTATAACCTACATCAACTAGGCTCGAGCAAGGCCGGAAGTGGGTT	1140
Qy		
1081	GACAAGCCGGGGCTTATAACCTACATCAACTAGGCTCGAGCAAGGCCGGAAGTGGGTT	1140
Db		
1141	TATGTGCCGAAGTTCGGGGCGGGTGCATCCACGAGTATACTGGTAACCTCGCAGGCTGG	1200
Qy		
1141	TATGTGCCGAAGTTCGGGGCGGGTGCATCCACGAGTATACTGGTAACCTCGCAGGCTGG	1200
Db		
1201	GTAGACAAGTACGCTTACTCAAGCGGTGGGTCTATCTCGAAGCTCAGGTTACGACCT	1260
Qy		
1201	GTAGACAAGTACGCTTACTCAAGCGGTGGGTCTATCTCGAAGCTCAGGTTACGACCT	1260
Db		
1261	GCCAAAGGGCAGTATGGCTACTCCGTGTGGAGCTACTCGGGGGTGGGCTGA	1311
Qy		
1261	GCCAAAGGGCAGTATGGCTACTCCGTGTGGAGCTACTCGGGGGTGGGCTGA	1311
Db		

## RESULT 5

```

US-10-385-305-5
/ Sequence 5, Application US/1039S305
/ Publication No. US20040018607A1
/ GENERAL INFORMATION:
/ APPLICANT: Callen, Walter
/ APPLICANT: Richardson, Toby
/ APPLICANT: Frey, Gerhard
/ APPLICANT: Short, Jay M.
/ APPLICANT: Mathur, Eric J.
/ APPLICANT: Gray, Kevin A.
/ APPLICANT: Kerovuo, Janne S.
/ APPLICANT: Slupska, Malgorzata
/ TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
/ TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
/ FILE REFERENCE: 09010-108001
/ CURRENT APPLICATION NUMBER: US/10/395,305
/ CURRENT FILING DATE: 2003-03-06
/ PRIOR APPLICATION NUMBER: US/10/081,872
/ PRIOR FILING DATE: 2002-02-21
/ PRIOR APPLICATION NUMBER: US 60/270,495
/ PRIOR FILING DATE: 2001-02-21
/ PRIOR APPLICATION NUMBER: US 60/270,496
/ PRIOR FILING DATE: 2001-02-21
/ PRIOR APPLICATION NUMBER: US 60/291,122
/ PRIOR FILING DATE: 2001-05-14
/ NUMBER OF SEQ ID NOS: 321
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 1311
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: synthetically generated oligonucleotides
US-10-385-305-5

```

Query Match 94.5%; Score 1239; DB 16; Length 1311;

QY	1	ATGCCCAAGTACTCCGAGCTGCGAAAGCGGGGTCTAATGACAGCGCTTCTACTCGGAC	60
Db	1	ATGCCCAAGTACTCCGAGCTGCGAAGCGGGCGTCTAATGACAGCGCTTCTACTCGGAC	60
QY	61	GTGGCTTCAGAGGAATAATGTTGGGACACATATCCGCAGGAAGATACCGAGTGGTACGAT	120
Ddb	61	GTCCAGGTGGAGGAATCTGTGTGGGACACCATCAGGAGCAAGATACCGAGTGGTACGAG	120
QY	121	GCGGGAATCTCCGCAATATGGAATCCCGCGGAGCAAGGGCATGGCGCGCGCTATTCCG	180

QY 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTCGCGGGTGGGCTGA 1311  
 DB 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTCGCGGGTGGGCTGA 1311

## RESULT 6

US-10-081-872-71  
 ; Sequence 71, Application US/10081872  
 ; Publication No. US20030125534A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT APPLICATION NUMBER: US/10/081.872  
 ; CURRENT FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 71  
 ; LENGTH: 1311  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated oligonucleotide  
 US-10-081-872-71

Query Match 93.9%; Score 1231; DB 15; Length 1311;  
 Best Local Similarity 96.2%; Pred. No. 0;  
 Matches 1261; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 ATGGCCAAAGTACTCCGAGCTGGAAGAGGGGGGTCAATATGACGGCTTCTACTGGGAC 60  
 DB 1 ATGGCCAAAGTACTCCGAGCTGGAAGAGGGGGGTCAATATGACGGCTTCTACTGGGAC 60

QY 61 GTGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120  
 DB 61 GTGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120

QY 121 GCCGGAATCTCCGCAATATGATATCCCGCGGAGCAAGGCAATGGCGCGCCCTATTTCG 180  
 DB 121 GCCGGAATCTCCGCAATATGATATCCCGCGGAGCAAGGCAATGGCGCGCCCTATTTCG 180

QY 181 ATGGGTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACAGAGGAGGACGTA 240  
 DB 181 ATGGGTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACAGAGGAGGACGTA 240

QY 241 GAGACGGCTTTGGCTCCAGCAGAGAGCTCGTGAAATGATATAACACCGCCACGCGCTAT 300  
 DB 241 GAGACGGCTTTGGCTCCAGCAGAGAGCTCGTGAAATGATATAACACCGCCACGCGCTAT 300

QY 301 GCGATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGCTGACCTGGAGTGG 360  
 DB 301 GCGATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGCTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTAACGACTATACCTGACCGCACTTCTCAAAAGTCCGCTCGGGTAATAC 420  
 DB 361 AACCCCTTCGTAACGACTATACCTGACCGCACTTCTCAAAAGTCCGCTCGGGTAATAC 420

QY 421 ACGGCCAACCTACCTCGACTTCCACCCGACGAGCTCCATGTCGGGGGATTCGCGAATTT 480

DB 421 ACTGCCAACTACCTCGACTTCCACCGACGAGCTCCATCGCGGGCATTCGGAACATTT 480  
 QY 481 GGAGGCTATCCGACATATGCCAGACGACAGAGCTGGGACAGTACTGGCTCTGGGCGCAGC 540  
 DB 481 GGAGGCTATCCGACATATGCCAGACGACAGAGCTGGGACAGTACTGGCTCTGGGCGCAGC 540

QY 541 CAGGAGAGCTACCGGCAATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
 DB 541 CAGGAGAGCTACCGGCAATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600

QY 601 GTCAAGGGCTATGCTCCCTGGGTCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGG 660  
 DB 601 GTCAAGGGCTATGCTCCCTGGGTCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGG 660

QY 661 GTTGAGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCATATCTCGAGCGGT 720  
 DB 661 GTTGAGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCATATCTCGAGCGGT 720

QY 721 GCCAAGSTCTTTGACTTGGCCCTCTACTACAAGATGATGAGGCTTTGACAAACAAAC 780  
 DB 721 GCCAAGSTCTTTGACTTGGCCCTCTACTACAAGATGATGAGGCTTTGACAAACAAAC 780

QY 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAACGGCCAGACTGTTGTCTCCCGCGACCCGCTTC 840  
 DB 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAACGGCCAGACTGTTGTCTCCCGCGACCCGCTTC 840

QY 841 AAGCCGCTAACCTTTGTAGCAACACGACACCGATATATCTTGGAACAAAGTATCCAGCC 900  
 DB 841 AAGCCGCTAACCTTTGTAGCAACACGACACCGATATATCTTGGAACAAAGTATCCAGCC 900

QY 901 TAGCGGTTTCATCTCTACCTACGAGGCGCAGCGCAATAATCTTACCGGACTTACGAGGAG 960  
 DB 901 TAGCGGTTTCATCTCTACCTACGAGGCGCAGCGCAATAATCTTACCGGACTTACGAGGAG 960

QY 961 TGGCTCAACAAAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTCGCCCGAGGA 1020  
 DB 961 TGGCTCAACAAAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTCGCCCGAGGA 1020

QY 1021 AGCACCGACATAGTCTACTACGATACGATGAACCTCATCTCGTCAGGAACGGCTACCGG 1080  
 DB 1021 AGCACCGACATAGTCTACTACGATACGATGAACCTCATCTCGTCAGGAACGGCTACCGG 1080

QY 1081 GACAAAGCGGGCTTATAACCTTACATCAACCTTAGGCTCGAGCAAGCGCGGAGTGGGTT 1140  
 DB 1081 GACAAAGCGGGCTTATAACCTTACATCAACCTTAGGCTCGAGCAAGCGCGGAGTGGGTT 1140

QY 1141 TATGTGCCGAAGTTCGCGGGCGGTGCATCCACGAGTATCTGGTAACTCGGAGGCTGG 1200  
 DB 1141 TATGTGCCGAAGTTCGCGGGCGGTGCATCCACGAGTATCTGGTAACTCGGAGGCTGG 1200

QY 1201 GTAGACAAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 DB 1201 GTAGACAAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

QY 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTCGCGGGTGGGCTGA 1311  
 DB 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTCGCGGGTGGGCTGA 1311

## RESULT 7

US-10-385-305-71  
 ; Sequence 71, Application US/10385305  
 ; Publication No. US20040018607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata



; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated oligonucleotide
US-10-385-305-71

Query Match 93.9%; Score 1231; DB 16; Length 1311;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1261; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 ATGGCCAGTACTCGAGTGGAAAGCGGGGTCTAATATCGAGGGTCTTACTGGGAC 60
DB 1 ATGGCCAGTACTCGAGTGGAAAGCGGGGTCTAATATCGAGGGTCTTACTGGGAC 60
QY 61 GTGGCTTCAGGAGGAATATGGTGGGACACAATACCGGAGAGATACCGAGGTGGTACGAT 120
DB 61 GTGGCTTCAGGAGGAATATGGTGGGACACAATACCGGAGAGATACCGAGGTGGTACGAT 120
QY 121 GCCGGAATCTCCGGAATATGGATCCCGCGGAGCAAGGGCATGGGCGGCCCTATTTCG 180
DB 121 GCCGGAATCTCCGGAATATGGATCCCGCGGAGCAAGGGCATGGGCGGCCCTATTTCG 180
QY 181 ATGGGCTACGACCCCTACGACTTCTTACCTCGGTAGTACGACCAAGGAGACGTA 240
DB 181 ATGGGCTACGACCCCTACGACTTCTTACCTCGGTAGTACGACCAAGGAGACGTA 240
QY 241 GAGACGCGTTTGGTCCAGCAGGAGCTCGTGAACATGATAAACCACCGCCACGCTAT 300
DB 241 GAGACGCGTTTGGTCCAGCAGGAGCTCGTGAACATGATAAACCACCGCCACGCTAT 300
QY 301 GCATGAGGTAATAGCGGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
DB 301 GCATGAGGTAATAGCGGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
QY 361 AACCCCTTCGTGAACGACTATACCTTGGACCGACTTCTCAAAAGTTCGGTAAATAC 420
DB 361 AACCCCTTCGTGAACGACTATACCTTGGACCGACTTCTCAAAAGTTCGGTAAATAC 420
QY 421 AGGCGCAACTACTCGACTTCCACCGGAGCTCCATGCGGCGGATTCGGAACATT 480
DB 421 AGGCGCAACTACTCGACTTCCACCGGAGCTCCATGCGGCGGATTCGGAACATT 480
QY 481 GGAGGCTATCCGACATATCCGACAGAGAGCTGGGACGAGTACTGGCTTGGGCGAGC 540
DB 481 GGAGGCTATCCGACATATCCGACAGAGAGCTGGGACGAGTACTGGCTTGGGCGAGC 540
QY 541 CAGGAGAGCTACGGGCGCATATCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGATAC 600
DB 541 CAGGAGAGCTACGGGCGCATATCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGATAC 600
QY 601 GTCAGGCTATGTCCTCGGTGCTCAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
DB 601 GTCAGGCTATGTCCTCGGTGCTCAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
QY 661 GTTGAGAGTACTGGGACACCAAGTTCGACGCTTCTCAACTGGGCATACCTGAGCGGT 720

DB 661 GTTGAGAGTACTGGGACACCAAGTTCGACGCTTCTCAACTGGGCATACCTGAGCGGT 720
QY 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGATGAGGCTTTGACAAACAAAAC 780
DB 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGATGATGAGGCTTTGACAAACAAAAC 780
QY 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAAAGCGGCAGACTGTTGCTCCCGGACCCGTT 840
DB 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAAAGCGGCAGACTGTTGCTCCCGGACCCGTT 840
QY 841 AAGCGCTAAGCTTTGTAGCAACACCGACCGATATAATCTGGAACAAAGTATCCAGCC 900
DB 841 AAGCGCTAAGCTTTGTAGCAACACCGACCGATATAATCTGGAACAAAGTATCCAGCC 900
QY 901 TACCGCTTCATCTCCTACCTACGAGGCGCAGCAGATATTTCTACCGGACTACGAGG 960
DB 901 TACCGCTTCATCTCCTACCTACGAGGCGCAGCAGATATTTCTACCGGACTACGAGG 960
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTTCGCCGAGGA 1020
DB 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTTCGCCGAGGA 1020
QY 1021 AGCACCAGCATAGTCTACTAGATTAAGTCACTCACTAGGCTTCGAGCAAGCGGAGTGG 1080
DB 1021 AGCACCAGCATAGTCTACTAGATTAAGTCACTCACTAGGCTTCGAGCAAGCGGAGTGG 1080
QY 1081 GACAAGCGCGGCTTATAAATCACTCACTAGGCTTCGAGCAAGCGGAGTGGGTT 1140
DB 1081 GACAAGCGCGGCTTATAAATCACTCACTAGGCTTCGAGCAAGCGGAGTGGGTT 1140
QY 1141 TATGTCGCGAAGTTCGCGGCGGTGATCCAGAGTATCTGTAACCTTCGAGGAGCTGG 1200
DB 1141 TATGTCGCGAAGTTCGCGGCGGTGATCCAGAGTATCTGTAACCTTCGAGGAGCTGG 1200
QY 1201 GTAGACAGTACGCTTACTCAAGCGGTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
DB 1201 GTAGACAGTACGCTTACTCAAGCGGTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
QY 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGAGACTACTGCGGGTGGGCTGA 1311
DB 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGAGACTACTGCGGGTGGGCTGA 1311

RESULT 8
US-10-081-872-37
; Sequence 37, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerovuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence

FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-37

Query Match 91.0%; Score 1192.6; DB 15; Length 1311;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 1237; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATGGCCAGTACTCCGAGCTGGAAGGCGGGGTCTAATGCAAGCGTTCTACTGGAC 60  
DB 1 ATGGCCAAATGCTCGAGCTCGAAGAGGGCGGGTCTAATGCAAGCGTTCTACTGGAC 60

QY 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120  
DB 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120

QY 121 GCCGAAATCTCGCAATATGATTTCCCGCGGAGCAAGGGCATGGGCGGCTATTGCG 180  
DB 121 GCCGAAATCTCGCAATATGATTTCCCGCGGAGCAAGGGCATGGGCGGCTATTGCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAACGGTA 240  
DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAACGGTA 240

QY 241 GAGAGCGCTTTGGCTCAACGAGAGCTCGTGAACATGATTAACACCGCCACGGCTAT 300  
DB 241 GAGAGCGCTTTGGCTCAACGAGAGCTCGTGAACATGATTAACACCGCCACGGCTAT 300

QY 301 GCGATGAAGGTAATAGCGATATAGTCAATCAACCGCGCGCGGTGACCTGGAGTGG 360  
DB 301 GCGATGAAGGTAATAGCGATATAGTCAATCAACCGCGCGCGGTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTGACGACTATCTGACGAGCTTCTCAAGTTCGCGTGGGTAATATAC 420  
DB 361 AACCCCTTCGTGACGACTATCTGACGAGCTTCTCAAGTTCGCGTGGGTAATATAC 420

QY 421 ACGGCAACTACCTCGACTTCCACCCGAAACGAGCTCCATCGCGCGGATTCGGGAACATTT 480  
DB 421 ACGGCAACTACCTCGACTTCCACCCGAAACGAGCTCCATCGCGCGGATTCGGGAACATTT 480

QY 481 GGAGGCTATCCGACATATGCAAGCAAGAGCTGGGACCACTGCTGGGCGAGC 540  
DB 481 GGAGGCTATCCGACATATGCAAGCAAGAGCTGGGACCACTGCTGGGCGAGC 540

QY 541 CAGGAGAGCTACGGGCAATATCTCAGGAGCATCGGATCGATCGTGGCGTTCGACTAC 600  
DB 541 CAGGAGAGCTACGGGCAATATCTCAGGAGCATCGGATCGATCGTGGCGTTCGACTAC 600

QY 601 GTCAGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAATCTGGTGGGAGGCTGGGG 660  
DB 601 GTCAGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAATCTGGTGGGAGGCTGGGG 660

QY 661 GTTGAGAGTACTGGGACACCAAGCTCGAGCTGTTCTCAACTGGGCATCTCGAGCGGT 720  
DB 661 GTTGAGAGTACTGGGACACCAAGCTCGAGCTGTTCTCAACTGGGCATCTCGAGCGGT 720

QY 721 GCCAAGTCTTGACTTCGCGCTCTACTCAAGATGATGAGCGCTTTGACACAAAC 780  
DB 721 GCCAAGTCTTGACTTCGCGCTCTACTCAAGATGATGAGCGCTTTGACACAAAC 780

QY 781 ATTCAGGCTCGCTCTCGCGCTTCAGACGGGCAAGCTGTTCTTCGCGGACCGCTTC 840  
DB 781 ATTCAGGCTCGCTCTCGCGCTTCAGACGGGCAAGCTGTTCTTCGCGGACCGCTTC 840

QY 841 AAGGCGTAACCTTTGTAGCAAAACCAAGCAACCGATATATCTGGAACAAAGTATCAGCC 900  
DB 841 AAGGCGTAACCTTTGTAGCAAAACCAAGCAACCGATATATCTGGAACAAAGTATCAGCC 900

QY 901 TACGCTTCATCTCACTACGAGGGCGGCGACCAATATTTCTACCGGCTACGAGGAG 960  
DB 901 TACGCTTCATCTCACTACGAGGGCGGCGACCAATATTTCTACCGGCTACGAGGAG 960

QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020

DB 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTCCCGAGGA 1020  
QY 1021 AGCACCGACATAGTCTACTACGATAACGATGAACACTCATCTTCTCAGGAACGGCTACGGG 1080  
DB 1021 AGCACTGACATCGTTTACTACGACACGACGAGCTGATATTCTGAGAAACGGCTACGGA 1080  
QY 1081 GACAAGCGGGGCTTATAACCTACATCACTAGGCTCGAGCAAGGCGGGAAGGTGGGTT 1140  
DB 1081 AGCAAGCGGGGCTTATAACCTACATCACTAGGCTCGAGCAAGGCGGGAAGGTGGGTT 1140  
QY 1141 TATGTCCGAAGTTCGGGGCGGCTGCATCCAGAGTATCTGGAACCTCGAGGCTGG 1200  
DB 1141 TACGTTCCGAAGTTCGGGGCGGCTGCATCCAGAGTATCTGGAACCTCGAGGCTGG 1200  
QY 1201 GTAGACAGTACGCTACTCAAGCGGCTGGGCTATCTCGAAGCTCCAGCTTACGACCT 1260  
DB 1201 GTGACAAAGTGGGTGGACTCAAGCGGCTGGGCTTACCTCGAGGCTCTCTCCACGACCG 1260  
QY 1261 GCCAACGGGAGTATGCTACTCCGTGTGGAGTACTCGCGGCTGGGCTGA 1311  
DB 1261 GCCAACGGGAGTATGCTACTCCGTGTGGAGTACTCGCGGCTGGGCTGA 1311

## RESULT 9

US-10-385-305-37  
; Sequence 37, Application US/10385305  
; Publication No. US20040018607A1  
GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupski, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/385,305  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-385-305-37

Query Match 91.0%; Score 1192.6; DB 16; Length 1311;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 1237; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATGGCCAGTACTCCGAGCTGGAAGGCGGGGTCTAATGCAAGCGTTCTACTGGAC 60  
DB 1 ATGGCCAAATGCTCGAGCTCGAAGAGGGCGGGTCTAATGCAAGCGTTCTACTGGAC 60

QY 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACCGCAGAGATACCGGAGTGGTACGAT 120  
DB 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACCGCAGAGATACCGGAGTGGTACGAT 120

QY 121 GCCGAAATCTCGCAATATGATTTCCCGCGGAGCAAGGGCATGGGCGGCTATTGCG 180

Db 121 GCCGGAATCTCCGCAATATGATATCCCGCGGAGCAAGGGCATGGCGCGCTATTTCG 180  
Qy 181 ATGGGCTACAGCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Db 181 ATGGGCTACAGCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTTCAAGCAGGAGCTCGTGAAACATGATATAAACACCCGCCAOCGCTAT 300  
Db 241 GAGACGGCTTTGGCTTCAAGCAGGAGCTCGTGAAACATGATATAAACACCCGCCAOCGCTAT 300  
Qy 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTGTGAACGACTATACCTGGACCGACTTCTCAAGGTGCGTTCGGGTAAATAC 420  
Db 361 AACCCCTTGTGAACGACTATACCTGGACCGACTTCTCAAGGTGCGTTCGGGTAAATAC 420  
Qy 421 ACGGCCAACTACCTCGACTTCCACCGAGAGCTCCATGCGGGGATTCGGAAATTT 480  
Db 421 ACGGCCAACTACCTCGACTTCCACCGAGAGCTCCATGCGGGGATTCGGAAATTT 480  
Qy 481 GAGGGTATCCGACATATGCCACGAAGAGCTGGACCACTACTGGCTCTGGGCCAGC 540  
Db 481 GAGGGTATCCGACATATGCCACGAAGAGCTGGACCACTACTGGCTCTGGGCCAGC 540  
Qy 541 CAGGAGAGTACGGCGCATATCTCAGGAGATCGGATCGCTTGGCGCTTCGACTAC 600  
Db 541 CAGGAGAGTACGGCGCATATCTCAGGAGATCGGATCGCTTGGCGCTTCGACTAC 600  
Qy 601 GTCAAGGGCTATGCTCCCTCGGTCTCAAGGAGCTGGTGAAGTGTGGGGAGCTGGCG 660  
Db 601 GTCAAGGGCTATGCTCCCTCGGTCTCAAGGAGCTGGTGAAGTGTGGGGAGCTGGCG 660  
Qy 661 GTTGAGAGTACTGGGACCAACCGTGCAGCGTGTCTCAACTGGGCATCTCGAGCGGT 720  
Db 661 GTTGAGAGTACTGGGACCAACCGTGCAGCGTGTCTCAACTGGGCATCTCGAGCGGT 720  
Qy 721 GCCAAGGCTTTGACTTCGCTCTACTACAGATGAGTGGGCTTTGACAAACAAAC 780  
Db 721 GCCAAGGCTTTGACTTCGCTCTACTACAGATGAGTGGGCTTTGACAAACAAAC 780  
Qy 781 ATTCAGGCTGTCTGCTGCTTCCAGACGCGCAGACTGTGTCTCCCGGACCGGTT 840  
Db 781 ATTCAGGCTGTCTGCTGCTTCCAGACGCGCAGACTGTGTCTCCCGGACCGGTT 840  
Qy 841 AAGGCGTAACCTTTGTAGCAACCAACGACCGCATATATCTGGAAACAAGTATCCAGCC 900  
Db 841 AAGGCGTAACCTTTGTAGCAACCAACGACCGCATATATCTGGAAACAAGTATCCAGCC 900  
Qy 901 TACGGCTTATCCTCACCTACGAGGCGCAGCGCAATATTTCTACCGGACTACGAGGAG 960  
Db 901 TACGGCTTATCCTCACCTACGAGGCGCAGCGCAATATTTCTACCGGACTACGAGGAG 960  
Qy 961 TGGCTCAACAGGATAAGCTCAAGAACTCTATCTGGATACATGAGAACTCCCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAAGCTCAAGAACTCTATCTGGATACATGAGAACTCCCGGAGGA 1020  
Qy 1021 AGCAGGACATAGTCTACTACGATAACGATGAACTCATCTTCGTCAGGAACGGTACGG 1080  
Db 1021 AGCAGGACATAGTCTACTACGATAACGATGAACTCATCTTCGTCAGGAACGGTACGG 1080  
Qy 1081 GACAGCGGGGCTTATAACCTCATCACTAGGCTCGAGCAAGGCGGAAAGTGGGTT 1140  
Db 1081 GACAGCGGGGCTTATAACCTCATCACTAGGCTCGAGCAAGGCGGAAAGTGGGTT 1140  
Qy 1141 TATGTCCGGAAGTTTCGGGGCGGTGATCCACAGATATAGTGAATCTCGGAGGCTGG 1200  
Db 1141 TATGTCCGGAAGTTTCGGGGCGGTGATCCACAGATATAGTGAATCTCGGAGGCTGG 1200  
Qy 1201 GTACAGAGTACGCTACTCAAGCGGTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

Db 1201 GTGGACAAAGTGGGTGAGTCAAGCGGCTGCTACCTCGAGGCTCTCTCCACGACCCG 1260  
Qy 1261 GCCAACGGCAGTAGTGGCTACTCTCGTGTGGAGTACTCTCGGGGTGGGTGA 1311  
Db 1261 GCCAACGGCAGTAGTGGCTACTCTCGTGTGGAGTACTCTCGGGGTGGGTGA 1311  
RESULT 10  
US-10-081-872-39  
; Sequence 39, Application US/10081872  
; Publication No. US20030125534A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-39

Query Match 90.5%; Score 1186.2; DB 15; Length 1311;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 1233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
Qy 1 ATGGCAAGTACTCCGAGCTCGGAAAGGGCGGGTCAATATGAGCGGTCTTACTGGGAC 60  
Db 1 ATGGCAAGTACTCCGAGCTCGGAAAGGGCGGGTCAATATGAGCGGTCTTACTGGGAC 60  
Qy 61 GTCCCTTCAGGAGGATATGGTGGGACAAATACGACAGATACCGAGTGGTACGAT 120  
Db 61 GTCCCTTCAGGAGGATATGGTGGGACAAATACGACAGATACCGAGTGGTACGAT 120  
Qy 121 GCCGGAATCTCCGCAATATGGATTCTCCCGCGAGCAAGGATGGCGCGCTATTTCG 180  
Db 121 GCCGGAATCTCCGCAATATGGATTCTCCCGCGAGCAGGGGTATGAGCGCGCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACAGAGGAAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACAGAGGAAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTTCCCAAGCAGGAGCTCGTGAAACATGATATAAACACCCGCCAOCGCTAT 300  
Db 241 GAGACGGCTTTGGCTTCCCAAGCAGGAGCTCGTGAAACATGATATAAACACCCGCCAOCGCTAT 300  
Qy 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTGTGAACGACTATACCTGGACCGACTTCTCAAGGTGCGTTCGGGTAAATAC 420  
Db 361 AACCCCTTGTGAACGACTATACCTGGACCGACTTCTCAAGGTGCGTTCGGGTAAATAC 420

QY	421	ACGGCAACTACCTCGACTTCACCCGAAACGAGCTCCATGCGGGCGAATTCGGAAACAATTT	480
DB	421	ACGGCCAACTACCTCGACTTCACCCGAAACGAGCTCCATGCGGGCGAATTCGGAAACAATTT	480
QY	481	GGAGGCTATCCGACATATGCACCAAGAGCTGGGAACCACTACTGGCTCTCGGGCCAGC	540
DB	481	GGAGGCTATCCGACATATGCACCAAGAGCTGGGAACCACTACTGGCTCTCGGGCCAGC	540
QY	541	CAGGAGAGCTACGGGGCANATCTCAGAGCATCGCATCGATGCTGCGGCTTCGACTAC	600
DB	541	CAGGAGAGCTACGGGGCANATCTCAGAGCATCGGTATCGATGCTGCGGCTTCGACTAC	600
QY	601	GTCAAAGGCTATGCTCCCTGGGTGTCGAAGGACTGGCTGAACCTGGTGGGGAGGCTGGGG	660
DB	601	GTGAAGGCTACGGAGCGTGGGTGTCGAAGGACTGGCTCAACTGGTGGGGCGGCTGGGCC	660
QY	661	GTTGAGAGTACTGGGACACCAAGCTCGAGCTGTTCTCAACTGGGCACTACTCCAGCGGT	720
DB	661	GTTGCGAGTACTGGGACCCCAAGTTGATGCCCTCCTCCCTGGGCTACTCAGGCGGC	720
QY	721	GCCAAAGTCTTGACTTCGGCCCTACTACAAGATGATGAGGCGCTTTGACAAACAAAAC	780
DB	721	GCCAAAGTCTTCGACTTCCCGCTCTACTACAAGATGATGAGGCGCTTTGACAAACAAAAC	780
QY	781	ATTCCAGCGCTCGTCTCTGCCCTTCAGAACGGCCAGACTGTTGTCTCCCGGACCCGCTC	840
DB	781	ATTCCAGCGCTCGTCTCTGCCCTTCAGAACGGCCAGACTGTTGTCTCCCGGACCCGCTC	840
QY	841	AAGCCGCTAACTTTGTAGAACACACGACACCGATATTAATCTGAAACAAGTATCCAGCC	900
DB	841	AAGGCCGTAACCTTTGTAGCAACACGATACCGATATTAATCTGAAACAAGTATCCAGCC	900
QY	901	TACGCGTTCACTCTCACTACGAGGGCCAGCCGACAATAATCTACCGGACTACGAGGAG	960
DB	901	TACGGTTCACTCTCACTACGAGGGCCAGCCGACAATAATCTACCGGACTACGAGGAG	960
QY	961	TGGCTCAACAGGATAGCTCAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGA	1020
DB	961	TGGCTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGA	1020
QY	1021	AGCACCGCATAGTCTACTACGATAACGATGAACCTCATCTCGTCAGGAACGGCTACGGG	1080
DB	1021	AGCACCGCATAGTCTACTACGATAACGATGAACCTCATCTCGTCAGGAACGGCTACGGG	1080
QY	1081	GACAGCCGGGCTTAAACCTACATCAACCTAGGCTCGAGCAAGCCGGAAGTGGGTT	1140
DB	1081	GACAAGCCGGGCTTAAACCTACATCAACCTAGGCTCGAGCAAGCCGGAAGTGGGTT	1140
QY	1141	TATGTGCCCAAGTTTCGGGGCGGTGCATCCAGAGTATACTGTGTAACTCCGAGGCTGG	1200
DB	1141	TACGTTTCGGAAGTTTCGGGGAGGTGCAATCCAGGATACACCGCAACTCCGGGCGGTG	1200
QY	1201	GTAGACAGTACGCTACTCAAGGCGCTGGGTCTACTCGAAGCTCCAGCTTACGACCT	1260
DB	1201	GTGACAAAGTGGGTGAGCTCAAGCGGTGGGTGTACTCTGAGGCCCTTCCCGACGCCCG	1260
QY	1261	GCCAAAGGCTATGCTACTCCGTGTGAGACTACTCGGGGTGGGTGTA	1311
DB	1261	GCCAAAGGCTATTACGGCTACTCTGCTGAGCTACTCGGGGTGGGTGTA1311	

## RESULT 11

US-10-385-305-39

; Sequence 39, Application US/10385305

; Publication No. US20040018607A1

**GENERAL INFORMATION:**

; APPLICANT: Callen, Walter

; APPLICANT: Richardson, Toby

; APPLICANT: Frey, Gerhard

APPLICANT: Short, Jay M.  
APPLICANT: Mathews, Eric T.

APPLICANT: Mathur, Eric J.  
APPLICANT: Gray Kevin A

APPLICANT: GRAY, KEVIN A.  
APPLICANT: KEROMIO, JANNE S.

AFFILIATION: RETORVUO, UAHNE S

```

; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated oligonucleotide
US-10-385-305-39

```

Query Match	90.5%;	Score 1186.2;	DB 16;	Length 1311;
Best Local Similarity	94.1%;	Pred. No. 0;		
Matches 1233;	Conservative	0;	Mismatches 78;	Indels 0; Gaps 0
QY	1	ATGGCCAAAGTACTCCGAGCTGGAAAGGGGGGGTCTATAATGCAGGCGCTTCTACTGGGAC	60	
DB	1	ATGGCCAAAGTACTCCGAGCTCGAAGAGGGGGGGTCTATAATGCAGGCGTCTCTACTGGGAC	60	
QY	61	GTGCGCTTCAGGAGAAATATGGTGGGACAAATACGCAGAAATACCGAGTGGTACGAT	120	
DB	61	GTGCGCTTCAGGAGAAATATGGTGGGACACAAATACGCAGAAATACCGAGTGGTACGAT	120	
QY	121	GCGGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGGGGGGCGCTATTTCG	180	
DB	121	GCGGGAATCTCCGCAATATGGATTCTCCGCGAGCGGGGTATGAGCGGCGCTATTTCG	180	
QY	181	ATGGGCTACGACCCCTACGACTCTTTTGACCTCGGTAGTACGACAGAGGAAACGGTA	240	
DB	181	ATGGGCTACGACCCCTACGATTATTGTACCTCGGTAGTACTACGAGAGGAAACGGTG	240	
QY	241	GAGAGCGCTTTTGCTCCAGCAGGAGCTCGTGAACATGATAACACACGCCCCAGCCCTAT	300	
DB	241	GAACGAGGTTTCGGCTCAAGCAGGAGCTCTAAACATGATAAACACGCCCCAGCCCTAT	300	
QY	301	GGCATGAAGGTAATAGCCGATATAGTCATCAACCACCGCCCGCGGTGACTCGGAGTGG	360	
DB	301	GGCATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTCGGAGTGG	360	
QY	361	AACCCCTCGTACGACTACTCGACCGACTCTCTCAAGGTCCGTCGGTAAATAC	420	
DB	361	AACCCCTCGTGAAGCACTATACCTGACCGACTCTCTCAAGGTCCGTCGGTAAATAC	420	
QY	421	ACGGCCAACTACCTCGACTCTCCACCGAAACGAGCTCCATCGGGGCAATTCGGAACATT	480	
DB	421	ACGGCCAACTACCTCGACTTCCACCGAAACGAGCTCCATCGGGGCAATTCGGAACATT	480	
QY	481	GGAGCTATCCGACATATGCCACGACAGAGCTGGGACAGTACTGGCTCTGGGCCAGC	540	
DB	481	GGAGCTATCCGACATATGCCACGACAGAGCTGGGACAGTACTGGCTCTGGGCCAGC	540	
QY	541	CAGGAGACTACGGGGCATATCTCAGGAGCATCGGATCGATCGCTTCGCGCTTCGACTAC	600	
DB	541	CAGGAGACTACGGGGCATATCTCAGGAGCATCGGTATCGATCGCTTCGCGCTTCGACTAC	600	
QY	601	GTCAAGGCTATGCTCCCTGGTCTGAAGGACTGGCTGAACTGGTGGGAGGCTGGGGC	660	
DB	601	GTGAAGGCTACGAGAGGCTGGGTCTGAAGGACTGGCTCAACTGGTGGGCGGCTGGGCC	660	
QY	661	GTTGAGAGTACTGGGACACAAACGCTCGACGCTGTGTCTCAACTGGGCACTACTCGAGCGGT	720	

```
Db 661 GTTGGGAGTACTGGGACCCCAACGTTGATGCCCTCTCCCTGGGCTTACTCGAGCGC 720
Qy 721 GCCAAGGCTTTGACTTCCCTCTACTACAGATGGATGAGCCTTTGACACAAAC 780
Db 721 GCCAAGGCTTTGACTTCCCTCTACTACAGATGGATGAGCCTTTGACACAAAC 780
Qy 781 ATTCCAGCGCTGCTCTCTCCCTTACAGACGGCCAGACTGTGTCTCCCGGACCGTTC 840
Db 781 ATTCCAGCGCTGCTCTCTCCCTTACAGACGGCCAGACTGTGTCTCCCGGACCGTTC 840
Qy 841 AAGGCGGTAACTTTGAGCAACACGACGACGATATATCTGMAAAGTATCCAGCC 900
Db 841 AAGGCGGTAACTTTGAGCAACACGACGACGATATATCTGMAAAGTATCCAGCC 900
Qy 901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATCTTACCGGACTACGAGG 960
Db 901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATCTTACCGGACTACGAGG 960
Qy 961 TGGCTCAACAGGATAAGCTCAAGAACCTTCATCTGGATACATGACAACTCGCCGAGGA 1020
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTTCATCTGGATACATGACAACTCGCCGAGGA 1020
Qy 1021 AGCACCAGCATAGTCTACTACGATACGATGAACATCTTCTGTCAGGAAAGCTACGG 1080
Db 1021 AGCACCAGCATAGTCTACTACGATACGATGAACATCTTCTGTCAGGAAAGCTACGG 1080
Qy 1081 GACAGCCGGGCTTATACCTACATCAACCTTAGCTCGAGCAAGCCGGAAGTGGGT 1140
Db 1081 GACAGCCGGGCTTATACCTACATCAACCTTAGCTCGAGCAAGCCGGAAGTGGGT 1140
Qy 1141 TATGTCCGAAGTTCGGGCGCGGTCATCCACGAGTATAGTGTAACTCGAGGCTGG 1200
Db 1141 TATGTCCGAAGTTCGGGCGCGGTCATCCACGAGTATAGTGTAACTCGAGGCTGG 1200
Qy 1201 GTAGCAAGTACTGTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
Db 1201 GTAGCAAGTACTGTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
Qy 1261 GCCAAGGCGAGTATGGCTACTTCCGTGTGGAGCTACTCGGGGTGGGTGA 1311
Db 1261 GCCAAGGCGAGTATGGCTACTTCCGTGTGGAGCTACTCGGGGTGGGTGA 1311

RESULT 12
US-10-081-739a-1
; Sequence 29, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Prev, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Slupsko, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081.872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 1311
; TYPE: DNA
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated oligonucleotide
US-10-081-872-29
```

```
Query Match 90.2%; Score 1183; DB 15; Length 1311;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
```

```
Qy 1 ATGCCAAGTACTCCGAGCTGGAAAAGGCGGGGTCTAATATGACAGGCTTCTACTGGGAC 60
Db 1 ATGCCAAGTACTCCGAGCTCGAAGAGGCGGGGTCTAATATGACAGGCTTCTACTGGGAC 60
Qy 61 GTGGCTTTCAGAGGAATATGTGTGGACACAAATACGGCAGAGATACCGGAGTGTACGAT 120
Db 61 GTGGCTTTCAGAGGAATATGTGTGGACACAAATACGGTACGCCAGAGATACCGGAGTGT 120
Qy 121 GCCGGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGGCATGGGCGGCTTATTCG 180
Db 121 GCCGGAATTTCCGCAATATGATTTCCCGCGGAGCAAGGGCATGGGCGGCTTATTCG 180
Qy 181 ATGGGCTTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240
Db 181 ATGGGCTTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240
Qy 241 GAGACGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTAT 300
Db 241 GAGACGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTAT 300
Qy 301 GGCATGAGATATAGCCGATATAGTATCAACACCGCGCGGCGGTGACCTGGAGTGG 360
Db 301 GGCATGAGATATAGCCGATATAGTATCAACACCGCGCGGCGGTGACCTGGAGTGG 360
Qy 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGGGTAAATAC 420
Db 361 AACCCCTTCGTGGGACTACACTGGACCGACTTCTCAAAGGTTCGGGTAAATAT 420
Qy 421 ACGCCAATCTCTGCACTTCCACCGCAAGAGCTCCATCGCGGCGATTCGCGAAACATTT 480
Db 421 ACTGCCAATCTCTGCACTTCCACCGCAAGAGCTCCATCGCGGCGATTCGCGAAACATTT 480
Qy 481 GGAGGCTATCCGACATATGCCACGACAGAGCTGGGACGAGTACTGGCTCTGGGCGAG 540
Db 481 GGAGGCTATCCGACATATGCCACGACAGAGCTGGGACGAGTACTGGCTCTGGGCGAG 540
Qy 541 CAGGAGAGCTACGCGCATATCTCAGGAGCATCGGCATCGATCGCTGGGCTTCGACTAC 600
Db 541 CAGGAGAGCTACGCGCATATCTCAGGAGCATCGGCATCGATCGCTGGGCTTCGACTAC 600
Qy 601 GTCAAGGCTATGCTCCCTGGGTCTCAAGAGCTGCTGAACTGTGTGGGAGGCTGGGG 660
Db 601 GTCAAGGCTATGCTCCCTGGGTCTCAAGAGCTGCTGAACTGTGTGGGAGGCTGGGG 660
Qy 661 GTTGGAGAGTACTGGGACACCAACGTCGCGCTTCTCAACTGGGCGATCTCGAGCGGT 720
Db 661 GTTGGAGAGTACTGGGACACCAACGTCGCGCTTCTCAACTGGGCGATCTCGAGCGGT 720
Qy 721 GCCAAGGCTTTGACTTCCGCTCTACTACTAAGATGATGAGGCTTTGACAAACAAAC 780
Db 721 GCCAAGGCTTTGACTTCCGCTCTACTACTAAGATGATGAGGCTTTGACAAACAAAC 780
Qy 781 ATTCCAGCGCTGCTCTCTCCCTTACAGACGGCCAGACTGTGTCTCCCGGACCGTTC 840
Db 781 ATTCCAGCGCTGCTCTCTCCCTTACAGACGGCCAGACTGTGTCTCCCGGACCGTTC 840
Qy 841 AAGGCGGTAACTTTGAGCAACACGACGACGATATATCTGMAAAGTATCCAGCC 900
Db 841 AAGGCGGTAACTTTGAGCAACACGACGACGATATATCTGMAAAGTATCCAGCC 900
Qy 901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATCTTACCGGACTACGAGG 960
Db 901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATCTTACCGGACTACGAGG 960
```



Db 1201 GTAGCAAGTACCTCTACTCAAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGTACTTGGGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGTACTTGGGGGTGGGCTGA 1311

RESULT 14

US-10-081-872-13  
; Sequence 13, Application US/10081872  
; Publication No. US2003012534A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-13

Query Match 89.4%; Score 1171.8; DB 15; Length 1311;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1224; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
QY 1 ATGGCCAAAGTACTCGAGCTGMAAAGGCGGGGTCAATGCGAGGGCTTCTACTGGGAC 60  
Db 1 ATGGCCAAAGTACTCGAGCTGMAAAGGCGGGGTCAATGCGAGGGCTTCTACTGGGAC 60  
QY 61 GTGCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAGATACCGGAGTGGTACGAT 120  
Db 61 GTGCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAGATACCGGAGTGGTACGAT 120  
QY 121 GCCGGAATCTCGGCAATATGATATCCCGCGGAGCAAGGCGCGGCGGCGCTATTCG 180  
Db 121 GCCGGAATCTCGGCAATATGATATCCCGCGGAGCAAGGCGCGGCGGCGCTATTCG 180  
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTATGACCAAGAGGGAACGGTA 240  
QY 241 GAGACGGCTTTGGCTCAAGCAGAGCTGTGAAATGATATAACACCGGCCACGCTAT 300  
Db 241 GAGACGGCTTTGGCTCAAGCAGAGCTGTGAAATGATATAACACCGGCCACGCTAT 300  
QY 301 GGCATGAAGGTAATAGCGATATAGTATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTCATAGCGGACATCGTATAAACACCGCGCGGCGGAGACCTCGAGTGG 360  
QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGGGTAAATAC 420  
Db 361 AACCCCTTCGTGGGACTACACCTGGAGCGACTTCTCAAAGGTTCGGGTAAATAT 420

QY 421 ACGCCAACTACCTCGACTTCCACCCGAAACGAGCTCCATCGGGGCGATTCCGGAAACATTT 480  
Db 421 ACTCCAACTACCTCGACTTCCACCCGAAACGAGCTCAAGTGTGTGACGAGGCGACATTT 480  
QY 481 GGAGGCTATCCCGACATATGCCAGCAAAAGAGCTGGGACCAAGTACTGGCTCTTGGGCGAGC 540  
Db 481 GGAGGCTATCCCGACATATAGCCCAAGAGAGCTGGGACCAAGTACTGGCTCTTGGGCGAGC 540  
QY 541 CAGGAGAGCTACGGCGCATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC 600  
Db 541 GATGAGAGCTACGGCGCATATCTTAAGGAGCATCGGCGCTTATCGCTGGCGCTTCGACTAC 600  
QY 601 GTCGAAGGCTATGCTCCCTGGGTCTGTCGAAGGAGCTGGCTGAACCTGGTGGGAGGCTTGGGCG 660  
Db 601 GTCGAAGGCTACGGAGCGTGGGTCTGTCGAAGGAGCTGGCTGAACCTGGTGGGAGGCTTGGGCG 660  
QY 661 GTTGGAGAGTACTGGGACACCAAGGCTGAGCGCTGTTCTCAACTGGGCACTACTCGAGGGT 720  
Db 661 GTTGGGAGTACTGGGAGACCAAAACGTTGATGCACTGCTCAACTGGGCTACTCGAGCGAT 720  
QY 721 GCCAAGGTCTTTGACTTTCGCGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAAAAC 780  
Db 721 GCAAAAGTCTTCGACTTTCGCGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAAAAC 780  
QY 781 ATTCCAGCGCTCGTCTCTGCGCTTCAGAGCGGCGAGCTGTTCTCTCCGCGACCCGCTTC 840  
Db 781 ATTCCAGCGCTCGTCTCTGCGCTTCAGAGCGGCGAGCTGTTCTCTCCGCGACCCGCTTC 840  
QY 841 AAGGCGGTAACCTTTGTAGCAAAACCAAGCAGACGATATATCTGGAACAGTATCCAGGC 900  
Db 841 AAGGCGGTAACCTTTGTAGCAAAACCAAGCAGACGATATATCTGGAACAGTATCCAGGC 900  
QY 901 TAGCGGTTCACTCTACCTACGAGGCGGCGGACCAATATTTACCGCGACTACGAGGAG 960  
Db 901 TAGCGGTTCACTCTACCTACGAGGCGGCGGACCAATATTTACCGCGACTACGAGGAG 960  
QY 961 TGGCTCAACAGGATTAAGCTCAAGAACCTCACTGATACATGAGAACCTTCGCGCGAGGA 1020  
Db 961 TGGCTCAACAGGATTAAGCTCAAGAACCTCACTGATACATGAGAACCTTCGCGCGAGGA 1020  
QY 1021 AGCACGACATAGTCTACTACGATAACGATGAACCTCATCTTCTGTCAGGAAACGCTACGGG 1080  
Db 1021 AGCACTGACATAGTCTACTACGATAACGATGAACCTCATCTTCTGTCAGGAAACGCTACGGG 1080  
QY 1081 GACAAGCGGGGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140  
Db 1081 GACAAGCGGGGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140  
QY 1141 TATGTCCGAAAGTTCCGCGGCGGCTGCATCCACGAGTACTTGGTAAACCTCGGAGGCTCG 1200  
Db 1141 TATGTCCGAAAGTTCCGCGGCGGCTGCATCCACGAGTACTTGGTAAACCTCGGAGGCTCG 1200  
QY 1201 GTAGCAAGTACTGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGCAAGTACTGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGGCTGA 1311

RESULT 15  
US-10-385-305-13  
; Sequence 13, Application US/10385305  
; Publication No. US20040018607A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.



APPLICANT: Kerovuo, Janne S.  
APPLICANT: Slupska, Malgorzata  
TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
TITLE OF INVENTION: AND METHODS OF USE THEREOF  
FILE REFERENCE: 09010-108001  
CURRENT APPLICATION NUMBER: US/10/385,305  
CURRENT FILING DATE: 2003-03-06  
PRIOR APPLICATION NUMBER: US/10/081,872  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: US 60/270,495  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: US 60/270,496  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: US 60/291,122  
PRIOR FILING DATE: 2001-05-14  
NUMBER OF SEQ ID NOS: 321  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 1311  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetically generated oligonucleotide

US-10-385-305-13

Query Match 89.4%; Score 1171.8; DB 16; Length 1311;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1224; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 1 ATGCCAAGTACTCCGAGCTGGAAAAGGGCGGGTCAATAATCGAGCGTTCTACTGGGAC 60  
Db 1 ATGCCAAGTACTCCGAGCTGGAAAAGGGCGGGTCAATAATCGAGCGTTCTACTGGGAC 60

Qy 61 GTCCCTTCAGAGGAAATATGCTGGGACACAATACGCGCAGAGATACCGAGTGGTACGAT 120  
Db 61 GTCCCTTCAGAGGAAATATGCTGGGACACAATACGCGCAGAGATACCGAGTGGTACGAT 120

Qy 121 GCCGGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGGGATGCGCGGCTATTTCG 180  
Db 121 GCCGGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGGGATGCGCGGCTATTTCG 180

Qy 181 ATGGGCTACACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGCAACGGTA 240  
Db 181 ATGGGCTACACCCCTACGACTTCTTTGACCTCGGTGAGTATGACAGAGGCAACGGTA 240

Qy 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATMAACACCGCCACGCTAT 300  
Db 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATMAACACCGCAGCTAC 300

Qy 301 GGCATGAGGTATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAGGTATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGG 360

Qy 361 AACCCCTTCGTGACGACTATACCTGACCGACTTCTCAAAGTTCGCTCGGCTAAATAC 420  
Db 361 AACCCCTTCGTGAGGACTACACTGACCGACTTCTCAAAGTTCGCTCGGCTAAATAT 420

Qy 421 ACGGCCAATACCTTCGACTTCCACCCGAAAGAGTCCATCGCGGCGGATTCGGAACATT 480  
Db 421 ACTGCCAATACCTTCGACTTCCACCCCAAGAGGTCAAGTGTGTGACGAGGGGCAATT 480

Qy 481 GAGGCTATCCGACATATCCAGCAAGAGTGGGACGAGTGGCTCTGGCGCAGC 540  
Db 481 GAGGCTATCCGACATATGCGCCAGAGAGAGTGGGACGAGTGGCTCTGGCGCAGC 540

Qy 541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
Db 541 GATGAGAGCTACCGCGCTACCTAAGGAGCATCGCGCTGATGCTGGCGCTTCGACTAC 600

Qy 601 GTCAGGGCTATGCTCCCTGGTCTCAAGGAGTGGCTGAACTGGTGGGAGGCTGGGG 660  
Db 601 GTCAGGGCTACCGAGCGTGGTCTCAAGGAGTGGCTGAACTGGTGGGAGGCTGGGG 660

Search completed: June 29, 2004, 13:31:50  
Job time : 611 secs

Qy 661 GTTGGAGAGTACTGGGACACCAACGCTCGAGCTGTGTTCTCAACTGGGCATATCTGAGCGGT 720  
Db 661 GTCGGGAGTACTGGGACACCAACGTTGATGCACTGCTCACTGGGCTTACTCGAGCGAT 720

Qy 721 GCCAAGGTCTTTGACTTGGCCCTCTACTACAGATGGATGAGCCCTTTGACACAAAC 780  
Db 721 GCCAAGGTCTTTGACTTCCGCTCTACTACAGATGGATGAGCCCTTTGACACAAAC 780

Qy 781 ATTCCAGCGCTGCTCTGCTGCTTCAAGCGGCGAGACTGTTGTTCTCCCGCGACCCGTTTC 840  
Db 781 ATTCCAGCGCTGCTCTGCTGCTTCAAGCGGCGAGACTGTTGTTCTCCCGCGACCCGTTTC 840

Qy 841 AAGCCGTAACCTTTGTAGCAAAACCCAGCACCGCATATATCTGGAACAAAGTATCCAGCC 900  
Db 841 AAGCCGTAACCTTTGTAGCAAAACCCAGCACCGCATATATCTGGAACAAAGTATCCAGCC 900

Qy 901 TAGCGGTTTCATCTCTCACTACGAGGCGGCGGCGAGCAATATTTCTACCGCGACTACGAGGAG 960  
Db 901 TAGCGGTTTCATCTCTCACTACGAGGCGGCGGCGGCGAGCAATATTTCTACCGCGACTACGAGGAG 960

Qy 961 TGGCTCAACAAAGGATAGCTCAAGAACTCATCTGGATACATGAGAACTCTCGCGCGAGGA 1020  
Db 961 TGGCTCAACAAAGGATAGCTCAAGAACTCATCTGGATACATGAGAACTCTCGCGCGAGGA 1020

Qy 1021 AGCACCGCATAGTCTTACTAGGATGAGTCACTCTTCTGTCAGGAAACGCTACGGG 1080  
Db 1021 AGCACTGACATAGTCTTACTAGGATGAGTCACTCTTCTGTCAGGAAACGCTACGGG 1080

Qy 1081 GACAAAGCGGGGCTTATAAACCCTACATCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGTT 1140  
Db 1081 GACAAAGCGGGGCTTATAAACCCTACATCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGTT 1140

Qy 1141 TATGTGCCGAAGTTTCGGGGCGGTGTCATCCACGAGTATGTTGTTAACTCTCGAGGCTGG 1200  
Db 1141 TATGTGCCGAAGTTTCGGGGCGGTGTCATCCACGAGTATGTTGTTAACTCTCGAGGCTGG 1200

Qy 1201 GTAGACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGACAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

Qy 1261 GCCAAGCGGCGAGTATGCTACTCCGTGTGAGCTACTGCGGGGTGGCTGA 1311  
Db 1261 GCCAAGCGGCGAGTATGCTACTCCGTGTGAGCTACTGCGGGGTGGCTGA 1311



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:29:46 ; Search time 59 Seconds  
(without alignments)  
2087.978 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459  
Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVSWYSCGVG 436

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A: Genesep 29Jan04: \*  
1: Genesep1980s: \*  
2: Genesep1990s: \*  
3: Genesep2000s: \*  
4: Genesep2001s: \*  
5: Genesep2002s: \*  
6: Genesep2003as: \*  
7: Genesep2003bs: \*  
8: Genesep2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2459	100.0	436	6	ABG70531	Abg70531 Alpha amy
2	2459	100.0	436	6	ABP70154	Abp70154 Amino aci
3	2424	98.6	436	6	ABU03074	Abu03074 Alpha amy
4	2406	97.8	436	6	ABU03042	Abu03042 Alpha amy
5	2378	96.7	436	6	ABU03068	Abu03068 Alpha amy
6	2377	96.7	436	6	ABU03057	Abu03057 Alpha amy
7	2371	96.4	436	6	ABU03053	Abu03053 Alpha amy
8	2368	96.3	436	6	ABU03041	Abu03041 Alpha amy
9	2360	96.0	436	6	ABU03058	Abu03058 Alpha amy
10	2358	95.9	436	6	ABU03046	Abu03046 Alpha amy
11	2358	95.9	436	6	ABU03069	Abu03069 Alpha amy
12	2356	95.8	436	6	ABU03054	Abu03054 Alpha amy
13	2356	95.8	436	6	ABU03048	Abu03048 Alpha amy
14	2352	95.6	436	6	ABU03066	Abu03066 Alpha amy
15	2350	95.6	436	6	ABU03043	Abu03043 Alpha amy
16	2350	95.6	436	6	ABU03052	Abu03052 Alpha amy
17	2349	95.5	436	6	ABP96602	Abp96602 797GL3 al
18	2349	95.5	436	6	ABU03059	Abu03059 Alpha amy
19	2349	95.5	436	6	ABU03061	Abu03061 Alpha amy
20	2349	95.5	518	6	ABP96609	Abp96609 Plasmid p
21	2349	95.5	741	6	ABP96606	Abp96606 Maize 797
22	2349	95.5	820	6	ABP96610	Abp96610 Plasmid p
23	2347	95.4	436	6	ABU03045	Abu03045 Alpha amy
24	2344	95.3	454	6	ABP96607	Abp96607 Plasmid p
25	2344	95.3	460	6	ABP96622	Abp96622 797GL3 fu

26	2344	95.3	460	6	ABP96608	Abp96608 Plasmid p
27	2344	95.3	460	6	ABP96624	Abp96624 797GL3 fu
28	2342	95.2	436	6	ABU03051	Abu03051 Alpha amy
29	2338	95.1	436	6	ABU03044	Abu03044 Alpha amy
30	2338	95.1	436	6	ABU03050	Abu03050 Alpha amy
31	2334	94.9	436	6	ABU03064	Abu03064 Alpha amy
32	2334	94.9	436	6	ABU03071	Abu03071 Alpha amy
33	2332	94.8	436	6	ABU03040	Abu03040 Alpha amy
34	2331	94.8	436	6	ABU03049	Abu03049 Alpha amy
35	2331	94.8	436	6	ABU03056	Abu03056 Alpha amy
36	2327	94.6	436	6	ABU03070	Abu03070 Alpha amy
37	2321	94.4	436	6	ABU03055	Abu03055 Alpha amy
38	2314	94.1	432	6	ABG70536	Abg70536 Alpha amy
39	2314	94.1	432	6	ABU03076	Abu03076 Alpha amy
40	2311	94.0	432	6	ABU03077	Abu03077 Alpha amy
41	2298	93.5	436	6	ABU03065	Abu03065 Alpha amy
42	2298	93.5	461	6	ABG70537	Abg70537 Alpha amy
43	2298	93.5	461	6	ABU03078	Abu03078 Alpha amy
44	2290	93.1	436	6	ABU03060	Abu03060 Alpha amy
45	2283	92.8	436	6	ABU03062	Abu03062 Alpha amy

## ALIGNMENTS

### RESULT 1

ABG70531  
ID ABG70531 standard; protein; 436 AA.

AC ABG70531;

DT 20-JAN-2003 (first entry)

XX Alpha amylase polypeptide #1.

XX Alpha amylase; enzyme; starch liquefaction; starch linkage hydrolysis;  
KW starch breakdown catalysis; high-maltose syrup; high-glucose syrup; rice;  
KW germinated rice; corn; barley; wheat; legume; sweet potato; anti-staling;  
KW textile desizing; lignocellulosic fibre treatment; enzymatic de-linking;  
KW recycled paper pulp; alpha-glucosidic linkage; dough; crumb firming;  
KW crumb elasticity; palatability; improved flavour.

OS Unidentified.

XX WO200268597-A2.

XX 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005538.

XX 21-FEB-2001; 2001US-0270495P.

XX 21-FEB-2001; 2001US-0270496P.

XX 14-MAY-2001; 2001US-0291122P.

XX (DIVE-) DIVERSA CORP.

XX Callen W, Richardson T, Frey G, Miller C, Kazaoka M, Mathur EU;

XX Short JM;

XX WPI; 2003-018657/01.

XX N-PSDB; ABS55928.

XX Novel alpha amylase polypeptide, useful e.g. for starch liquefaction,  
PT textile desizing, treating lignocellulosic fibers and for preparing dough  
PT or baked product, has increased activity and stability.

XX Claim 30; Fig 7; 147pp; English.

XX The invention relates to a purified alpha amylase polypeptide and the  
CC polynucleotide encoding it. The polypeptide is useful for liquefying a  
CC starch containing composition, hydrolysing a starch linkage, catalysing  
CC the breakdown of a starch and producing high-maltose or high-glucose  
CC syrup or mixed syrup, where the starch is from rice, germinated rice,

corn, barley, wheat, legumes or sweet potato. The polypeptide is also useful for washing an object, textile desizing, treating lignocellulosic fibres for improving fibre properties, enzymatic de-inking of recycled paper pulp and modifying small molecules by biocatalytic reaction. The polypeptide is further useful for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, damaging fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore, by pumping an enzyme treatment comprising the polypeptide and an aqueous fluid to a desired location within the well-bore which is effective to attack the alpha-glucosidic linkages in the starch-containing fluid. The alpha amylases have increased activity and stability at increased pH and temperature and can be used in preparing dough to provide an improved anti-staling effect as measured by less crumb firming, retained crumb elasticity, improved slice-ability, improved palatability or improved flavour. This sequence represents an alpha amylase of the invention

XX SQ Sequence 436 AA;

Query Match 100.0%; Score 2459; DB 6; Length 436;  
Best Local Similarity 100.0%; Pred. No. 3.3e-203;  
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDYDAGISAIWIPPSKMGKGAYS 60  
DB 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDYDAGISAIWIPPSKMGKGAYS 60

QY 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLW 120  
DB 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLW 120

QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
DB 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180

QY 181 QESYAALYLRISIGIDAWRFDYKGYAPVVKDNLNMGWGAEGYWDTNDAVLNWAYSSG 240  
DB 181 QESYAALYLRISIGIDAWRFDYKGYAPVVKDNLNMGWGAEGYWDTNDAVLNWAYSSG 240

QY 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
DB 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300

QY 301 YAFILTYEGQPTIFPYRDEEELNKKLNLIIWHENLAGGSTDIVYDNDDELIFVRNGYG 360  
DB 301 YAFILTYEGQPTIFPYRDEEELNKKLNLIIWHENLAGGSTDIVYDNDDELIFVRNGYG 360

QY 361 DKPGLITVINLGSSKAGRWVVPKAGACIHEYTGNLGGWVDKYYSSGWVYLEAPAYDP 420  
DB 361 DKPGLITVINLGSSKAGRWVVPKAGACIHEYTGNLGGWVDKYYSSGWVYLEAPAYDP 420

QY 421 ANGOYGYSVMSYCGVG 436  
DB 421 ANGOYGYSVMSYCGVG 436

RESULT 2  
ABP70154  
ID ABP70154 standard; protein; 436 AA.  
AC ABP70154;  
XX  
XX 07-APR-2003 (first entry)  
DT  
XX Amino acid sequence of an alpha-amylase.  
XX Alpha-amylase; enzyme; detergent; baking; beverage; fuel; ethanol.  
XX Bacillus sp.  
OS  
XX  
XX WO200292802-A1.  
XX  
XX 21-NOV-2002.

14-MAY-2002; 2002WO-US015422.  
XX PF  
XX 14-MAY-2001; 2001US-0291122P.  
XX PR  
XX (DIVE-) DIVERSA CORP.  
XX PA  
XX Gerendash J;  
XX PI  
XX WPI; 2003-156728/15.  
XX DB N-FSDS; ABZ33371.  
XX DR  
XX Purifying enzyme, such as alpha-amylase and cellulase, involves  
PT flocculating fermentation broth containing bacterial cells containing the  
PT enzyme, centrifuging and extracting enzymes by filtration.  
XX  
XX Disclosure; Fig 7; 108pp; English.  
XX PS  
XX  
XX The present sequence represents an alpha-amylase enzyme, which may be  
CC purified using the method of the invention. The specification describes a  
CC method for purifying an enzyme, such as alpha-amylase and cellulase, by  
CC flocculating fermentation broth containing bacterial cells containing the  
CC enzyme, centrifuging and extracting enzymes by filtration. Alternatively,  
CC the above method involves subjecting a fermentation broth to a heat-  
CC killing procedure, where a resultant heat-killed broth is formed, washing  
CC the heat-killed broth with a buffered medium, releasing the enzyme  
CC contained in the cells of the fermentation broth and extracting the  
CC enzyme by microfiltration. The method is useful for purifying an enzyme  
CC derived from a mixed population of organisms or an isolate from the  
CC population. The alpha-amylases are useful in corn-wet milling processes,  
CC detergents, baking processes, beverages and in oil fields (fuel ethanol)  
XX SQ Sequence 436 AA;

Query Match 100.0%; Score 2459; DB 6; Length 436;  
Best Local Similarity 100.0%; Pred. No. 3.3e-203;  
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDYDAGISAIWIPPSKMGKGAYS 60  
DB 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDYDAGISAIWIPPSKMGKGAYS 60

QY 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLW 120  
DB 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLW 120

QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
DB 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180

QY 181 QESYAALYLRISIGIDAWRFDYKGYAPVVKDNLNMGWGAEGYWDTNDAVLNWAYSSG 240  
DB 181 QESYAALYLRISIGIDAWRFDYKGYAPVVKDNLNMGWGAEGYWDTNDAVLNWAYSSG 240

QY 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
DB 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300

QY 301 YAFILTYEGQPTIFPYRDEEELNKKLNLIIWHENLAGGSTDIVYDNDDELIFVRNGYG 360  
DB 301 YAFILTYEGQPTIFPYRDEEELNKKLNLIIWHENLAGGSTDIVYDNDDELIFVRNGYG 360

QY 361 DKPGLITVINLGSSKAGRWVVPKAGACIHEYTGNLGGWVDKYYSSGWVYLEAPAYDP 420  
DB 361 DKPGLITVINLGSSKAGRWVVPKAGACIHEYTGNLGGWVDKYYSSGWVYLEAPAYDP 420

QY 421 ANGOYGYSVMSYCGVG 436  
DB 421 ANGOYGYSVMSYCGVG 436

RESULT 3  
ABUC3074

ID ABU03074 standard; protein; 436 AA.  
 XX AC ABU03074;  
 XX DT 21-JAN-2003 (first entry)  
 XX DE Alpha amylase polypeptide #35.  
 XX KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 XX KW starch modification.  
 XX OS Unidentified.  
 XX PN WO200269589-A2.  
 XX PD 06-SEP-2002.  
 XX PF 21-FEB-2002; 2002WO-US005069.  
 XX PR 21-FEB-2001; 2001US-0270495P.  
 XX PR 21-FEB-2001; 2001US-0270496P.  
 XX PR 14-MAY-2001; 2001US-0291122P.  
 XX PA (DIVE-) DIVERSA CORP.  
 XX PI Callen W, Richardson T, Frey G;  
 XX DR WPI; 2003-018656/01.  
 XX DR N-PSDB; ABX08445.  
 XX PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 XX fibers, and for producing high-maltose or high-glucose syrup.  
 XX PS Claim 30; Fig 16; 301pp; English.  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 XX polypeptides of the invention  
 XX SQ Sequence 436 AA;  
 Query Match 98.6%; Score 2424; DB 6; Length 436;  
 Best Local Similarity 98.4%; Fred. No. 3.4e-200;  
 Matches 429; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIRQKIPEDYDAGISAIWIPSPAKSGMGAYS 60  
 DB 1 MAKYLEBEGGVIMQAFYWDVPSGGIWWDTIRQKIPEDYDAGISAIWIPSPAKSGMGAYS 60  
 QY 61 MGYPDPDFDLGVDQKQTVETREGSKOELVNMINTAHAYGMKIADIVINHRAGDLEW 120  
 DB 61 MGYPDPDFDLGVDQKQTVETREGSKOELVNMINTAHAYGIKVIADIVINHRAGDLEW 120

QY 121 NPFVNDYTDWTFDSKVAGSKYTANYLDFHPNHLHAGDSGTGGYPDICHKSDQYWLWAS 180  
 DB 121 NPFVGDVTDWTFDSKVAGSKYTANYLDFHPNHLHAGDSGTGGYPDICHKSDQYWLWAS 180  
 QY 181 QESYAAYLRSIGIDAWRFDYKGYAPWVVKDNLNMGCGWAVGEYWDTNVDVNLWAYSSG 240  
 DB 181 QESYAAYLRSIGIDAWRFDYKGYAPWVVKDNLNMGCGWAVGEYWDTNVDVNLWAYSSG 240  
 QY 241 AKVFDFALYYKMDFAFNKNI PALVSALONGQTWVS RDPKAVTFVANHTDIIWNKYPA 300  
 DB 241 AKVFDFALYYKMDFAFNKNI PALVSALONGQTWVS RDPKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGQPTTFYRDYEEWLNKDKLNIWIHENLAGGSTDIVYDNDDELIFVRNGY 360  
 DB 301 YAFILTYEGQPTTFYRDYEEWLNKDKLNIWIHENLAGGSTDIVYDNDDELIFVRNGY 360  
 QY 361 DKPLITVIINLGSSKAGRWVVPKFAACIHEVTGNLGGWVDKVVYSSGWVYLEAPAYDP 420  
 DB 361 SKPLITVIINLGSSKAGRWVVPKFAACIHEVTGNLGGWVDKVVYSSGWVYLEAPAYDP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 DB 421 ANGQYGSVWSYCGVG 436  
 RESULT 4  
 ABU03042  
 ID ABU03042 standard; protein; 436 AA.  
 XX AC ABU03042;  
 XX DT 21-JAN-2003 (first entry)  
 XX DE Alpha amylase polypeptide #3.  
 XX KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 XX KW starch modification.  
 XX OS Unidentified.  
 XX PN WO200268589-A2.  
 XX PD 06-SEP-2002.  
 XX PF 21-FEB-2002; 2002WO-US005068.  
 XX PR 21-FEB-2001; 2001US-0270495P.  
 XX PR 21-FEB-2001; 2001US-0270496P.  
 XX PR 14-MAY-2001; 2001US-0291122P.  
 XX PA (DIVE-) DIVERSA CORP.  
 XX PI Callen W, Richardson T, Frey G;  
 XX DR WPI; 2003-018656/01.  
 XX DR N-PSDB; ABX08445.  
 XX PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 XX fibers, and for producing high-maltose or high-glucose syrup.  
 XX PS Claim 30; Fig 16; 301pp; English.  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 XX polypeptides of the invention  
 XX SQ Sequence 436 AA;

CC	pulp, for producing a high-maltose or high-glucose syrup or a mixed	PD	06-SEP-2002.
CC	syrup, and for increasing the flow of production fluids from a	XX	
CC	subterranean formation by removing a viscous, starch-containing, damaging	PF	21-FEB-2002; 2002WO-US005068.
CC	fluid formed during production operations and found within the	XX	
CC	subterranean formation which surrounds a completed well bore. The	PR	21-FEB-2001; 2001US-0270495P.
CC	polypeptide is also useful for preparing a dough or a baked product	PR	21-FEB-2001; 2001US-0270496P.
CC	prepared from the dough and in corn-wet milling processes, detergents,	PR	14-MAY-2001; 2001US-0291122P.
CC	baking processes, beverages, oil fields (fuel ethanol), brewing processes	XX	
CC	and starch modification in the paper and pulp industry, for removing	PA	(DIVE-) DIVERSA CORP.
CC	starch containing stains from a material and for reducing staling of	PI	Callen W, Richardson T, Frey G;
CC	bakery products. Sequences ABU03040-ABU03144 represent alpha amylase	XX	
CC	polypeptides of the invention	XX	
XX		DR	WPI; 2003-018656/01.
SQ	Sequence 436 AA;	DR	N-PSDB; ABX08439.
		XX	
Query Match	97.8%; Score 2406; DB 6; Length 436;	PT	Novel purified polypeptide with alpha-amylase activity, useful e.g. for
Best Local Similarity	97.5%; Pred. No. 1.2e-198;	PT	liquefying starch, for textile desizing, for treating lignocellulosic
Matches 425; Conservative	6; Mismatches 5; Indels 0; Gaps 0;	PT	fibers, and for producing high-maltose or high-glucose syrup.
QY	1 MAKYSELEGGVIMQAFYNDVPSGGIWDITIRQKIPEDYDAGISAIWIPPSKMGGGAYS 60	XX	Claim 30; Fig 16; 301pp; English.
DB	1 MAKYSELEGGVIMQAFYNDVPSGGIWDITIRQKIPEDYDAGISAIWIPPSKMGGGAYS 60	XX	
QY	61 MGYPDYDFDLGEYDQGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120	CC	The invention relates to a purified polypeptide with alpha-amylase
DB	61 MGYPDYDFDLGEYDQGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120	CC	activity and the polynucleotide encoding it. The polypeptide is useful
QY	121 NPFVNDYTWDPSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180	CC	for hydrolysing starch linkages, for catalysing the breakdown of a
DB	121 NPFVNDYTWDPSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180	CC	starch, for modifying small molecules, for liquefying starch, for washing
QY	181 QESYAALRSIGIDANRFDYKGYAPWVKDNLNMGWAGVEYDNTVDVNLWAYSSG 240	CC	an object, for textile desizing, for treating lignocellulosic fibers, for
DB	181 QESYAALRSIGIDANRFDYKGYAPWVKDNLNMGWAGVEYDNTVDVNLWAYSSG 240	CC	improving fibre properties, for enzymatic de-inking of recycled paper
QY	241 AKVPDFALYKMDAEADFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300	CC	pulp, for producing a high-maltose or high-glucose syrup or a mixed
DB	241 AKVPDFALYKMDAEADFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300	CC	syrup, and for increasing the flow of production fluids from a
QY	301 YAFILTYEGQPTIFRYDYEELNKKDLNLIWHENLAGSGTDIVYDNDLIFVRNGYG 360	CC	subterranean formation by removing a viscous, starch-containing, damaging
DB	301 YAFILTYEGQPTIFRYDYEELNKKDLNLIWHENLAGSGTDIVYDNDLIFVRNGYG 360	CC	fluid formed during production operations and found within the
QY	421 ANGOYGYSVMSYCGVG 436	CC	subterranean formation which surrounds a completed well bore. The
DB	421 ANGOYGYSVMSYCGVG 436	CC	polypeptide is also useful for preparing a dough or a baked product
RESULT 5		CC	prepared from the dough and in corn-wet milling processes, detergents,
ABU03068		CC	baking processes, beverages, oil fields (fuel ethanol), brewing processes
ID	ABU03068 standard; protein; 436 AA.	CC	and starch modification in the paper and pulp industry, for removing
XX		CC	starch containing stains from a material and for reducing staling of
AC	ABU03068;	CC	bakery products. Sequences ABU03040-ABU03144 represent alpha amylase
XX		CC	polypeptides of the invention
DT	21-JAN-2003 (first entry)	XX	
XX		XX	Sequence 436 AA;
DE	Alpha amylase polypeptide #29.	XX	
XX		XX	Query Match
KW	Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;	XX	97.8%; Score 2406; DB 6; Length 436;
KW	starch breakdown catalysis; textile desizing; lignocellulosic fibre;	XX	Best Local Similarity
KW	enzymatic de-inking; recycled paper; high-maltose syrup; dough;	XX	97.5%; Pred. No. 1.2e-198;
KW	high glucose syrup; corn-wet milling process; detergent; baking process;	XX	Matches 425; Conservative
KW	beverage; oil field; fuel ethanol; brewing process; staling;	XX	6; Mismatches 5; Indels 0; Gaps 0;
KW	starch modification.	XX	
XX		XX	QY
OS	Unidentified.	XX	1 MAKYSELEGGVIMQAFYNDVPSGGIWDITIRQKIPEDYDAGISAIWIPPSKMGGGAYS 60
XX		DB	1 MAKYSELEGGVIMQAFYNDVPSGGIWDITIRQKIPEDYDAGISAIWIPPSKMGGGAYS 60
PN	WO200268589-A2.	QY	61 MGYPDYDFDLGEYDQGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120
XX		DB	61 MGYPDYDFDLGEYDQGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120
XX		QY	121 NPFVNDYTWDPSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
XX		DB	121 NPFVNDYTWDPSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
XX		QY	181 QESYAALRSIGIDANRFDYKGYAPWVKDNLNMGWAGVEYDNTVDVNLWAYSSG 240
XX		DB	181 QESYAALRSIGIDANRFDYKGYAPWVKDNLNMGWAGVEYDNTVDVNLWAYSSG 240
XX		QY	241 AKVPDFALYKMDAEADFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300
XX		DB	241 AKVPDFALYKMDAEADFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300
XX		QY	301 YAFILTYEGQPTIFRYDYEELNKKDLNLIWHENLAGSGTDIVYDNDLIFVRNGYG 360
XX		DB	301 YAFILTYEGQPTIFRYDYEELNKKDLNLIWHENLAGSGTDIVYDNDLIFVRNGYG 360
XX		QY	361 DKPGLITYINLGSKAGRWYVVPKFAACIHEYTGNLGGWVKRYVYSSGWVYLEAPAYDP 420
XX		DB	361 DKPGLITYINLGSKAGRWYVVPKFAACIHEYTGNLGGWVKRYVYSSGWVYLEAPAYDP 420
XX		QY	421 ANGOYGYSVMSYCGVG 436
XX		DB	421 ANGOYGYSVMSYCGVG 436
XX		XX	
XX		XX	Query Match
XX		XX	96.7%; Score 2378; DB 6; Length 436;
XX		XX	Best Local Similarity
XX		XX	96.1%; Pred. No. 3.1e-196;
XX		XX	Matches 419; Conservative
XX		XX	8; Mismatches 9; Indels 0; Gaps 0;
XX		XX	
XX		QY	1 MAKYSELEGGVIMQAFYNDVPSGGIWDITIRQKIPEDYDAGISAIWIPPSKMGGGAYS 60
XX		DB	1 MAKYSELEGGVIMQAFYNDVPSGGIWDITIRQKIPEDYDAGISAIWIPPSKMGGGAYS 60
XX		QY	61 MGYPDYDFDLGEYDQGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120
XX		DB	61 MGYPDYDFDLGEYDQGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120
XX		QY	121 NPFVNDYTWDPSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
XX		DB	121 NPFVNDYTWDPSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
XX		QY	181 QESYAALRSIGIDANRFDYKGYAPWVKDNLNMGWAGVEYDNTVDVNLWAYSSG 240
XX		DB	181 QESYAALRSIGIDANRFDYKGYAPWVKDNLNMGWAGVEYDNTVDVNLWAYSSG 240
XX		QY	241 AKVPDFALYKMDAEADFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300
XX		DB	241 AKVPDFALYKMDAEADFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300
XX		QY	301 YAFILTYEGQPTIFRYDYEELNKKDLNLIWHENLAGSGTDIVYDNDLIFVRNGYG 360
XX		DB	301 YAFILTYEGQPTIFRYDYEELNKKDLNLIWHENLAGSGTDIVYDNDLIFVRNGYG 360
XX		QY	361 DKPGLITYINLGSKAGRWYVVPKFAACIHEYTGNLGGWVKRYVYSSGWVYLEAPAYDP 420
XX		DB	361 DKPGLITYINLGSKAGRWYVVPKFAACIHEYTGNLGGWVKRYVYSSGWVYLEAPAYDP 420

QY	421	ANGQGYSVMSYCGVG	436
DB	421	ANGQGYSVMSYCGVG	436
RESULT 6			
ID	ABU03057	standard; protein; 436 AA.	
XX	AC	ABU03057;	
XX	AC	ABU03057;	
DT	21-JAN-2003	(first entry)	
XX	DE	Alpha amylase polypeptide #18.	
KW	KW	Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction; starch breakdown catalysis; textile desizing; lignocellulosic fibre; enzymatic de-linking; recycled paper; high-maltose syrup; dough; high glucose syrup; corn-wet milling process; detergent; baking process; beverage; oil field; fuel ethanol; brewing process; staling;	
OS	OS	Unidentified.	
PB	WO200268589-A2.		
PN	06-SEP-2002.		
PF	21-FEB-2002;	2002WO-US0005068.	
PR	21-FEB-2001;	2001US-0270495P.	
PR	21-FEB-2001;	2001US-0270496P.	
PR	14-MAY-2001;	2001US-0291122P.	
XX	PA	(DIVE-) DIVERSA CORP.	
PI	Callen W,	Richardson T, Frey G;	
DR	WI;	2003-018656/01.	
DR	N-PSDB;	ABX08428.	
XX	XX	Novel purified polypeptide with alpha-amylase activity, useful e.g. for liquefying starch, for textile desizing, for treating lignocellulosic fibers, and for producing high-maltose or high-glucose syrup.	
PT	Claim 30;	Fig 16; 30pp; English.	
XX	XX	The invention relates to a purified polypeptide with alpha-amylase activity and the polynucleotide encoding it. The polypeptide is useful for hydrolysing starch linkages, for catalysing the breakdown of a starch, for modifying small molecules, for liquefying starch, for washing an object, for textile desizing, for treating lignocellulosic fibers, for improving fibre properties, for enzymatic de-linking of recycled paper pulp, for producing a high-maltose or high-glucose syrup or a mixed syrup, and for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, damaging fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore. The polypeptide is also useful for preparing a dough or a baked product prepared from the dough and in corn-wet milling processes, detergents, baking processes, beverages, oil fields (fuel ethanol), brewing processes and starch modification in the paper and pulp industry, for removing starch containing stains from a material and for reducing staling of bakery products. Sequences ABU03040-ABU03144 represent alpha amylase polypeptides of the invention	
XX	SQ	Sequence 436 AA;	
Query Match 96.7%; Score 2377; DB 6; Length 436;			
Best Local Similarity 96.8%; Pred. No. 3.8e-196;			
Matches 422; Conservative 6; Mismatches 8; Indels 0; Gaps 0;			

high glucose syrup; corn-wet milling process; detergent; baking process; beverage; oil field; fuel ethanol; brewing process; stalling; starch modification.

Claim 30; Fig 16; 301pp; English.  
The invention relates to a purified polypeptide with alpha-amylase activity and the polynucleotide encoding it. The polypeptide is useful for hydrolysing starch linkages, for catalysing the breakdown of a starch, for modifying small molecules, for liquefying starch, for washing an object, for textile desizing, for treating lignocellulosic fibers, for improving fibre properties, for enzymatic de-inking of recycled paper pulp, for producing a high-maltose or high-glucose syrup or a mixed syrup, and for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, damaging fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore. The polypeptide is also useful for preparing a dough or a baked product prepared from the dough and in corn-wet milling processes, detergents, baking processes, beverages, oil fields (fuel ethanol), brewing processes and starch modification in the paper and pulp industry, for removing starch containing stains from a material and for reducing staling of bakery products. Sequences ABU03040-ABU03144 represent alpha amylase polypeptides of the invention

Novel purified polypeptide with alpha-amylase activity, useful e.g. for liquefying starch, for textile desizing, for treating lignocellulosic fibers, and for producing high-maltose or high-glucose syrup.  
Claim 30; Fig 16; 301pp; English.

Query Match 96.4%; Score 2371; DB 6; Length 436;  
Best Local Similarity 95.9%; Pred. No. 1.2e-195;  
Matches 418; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

1 MAKYSELEKGVIMQAFYWDVPSGGIWDITQKIPEDWDAGISAIWIPPSKMGCGAYS 60  
1 MAKYLEEGGLINQAFYWDVPMGGIWDITVAQKIPDWSAGISAIWIPPSKMGCGAYS 60  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
121 NPFVNDYTWTDFSKVSGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYPA 300  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYLA 300

1 MAKYSELEKGVIMQAFYWDVPSGGIWDITQKIPEDWDAGISAIWIPPSKMGCGAYS 60  
1 MAKYLEEGGLINQAFYWDVPMGGIWDITVAQKIPDWSAGISAIWIPPSKMGCGAYS 60  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
121 NPFVNDYTWTDFSKVSGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYPA 300  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYLA 300  
301 YAFILTYEGQPTIFPYRDEYEWANKKLNLIWIHLENLAGGSTDIVYDNDDELIFVRNGYG 360  
301 YAFILTYEGQPTIFPYRDEYEWANKKLNLIWIHLENLAGGSTDIVYDNDDELIFVRNGYG 360  
361 DKPGLIYINLSSKAGRWVYVPRFACAGTHEYTNLGNLWGWVYKVSNGWVYLEAPAYDP 420  
361 DKPGLIYINLSSKAGRWVYVPRFACAGTHEYTNLGNLWGWVYKVSNGWVYLEAPAYDP 420  
421 ANGQYGSVMSYCGVG 436  
421 ANGQYGSVMSYCGVG 436

Query Match 96.3%; Score 2368; DB 6; Length 436;  
Best Local Similarity 96.1%; Pred. No. 2.3e-195;  
Matches 419; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

1 MAKYSELEKGVIMQAFYWDVPSGGIWDITQKIPEDWDAGISAIWIPPSKMGCGAYS 60  
1 MAKYLEEGGLINQAFYWDVPMGGIWDITVAQKIPDWSAGISAIWIPPSKMGCGAYS 60  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
121 NPFVNDYTWTDFSKVSGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYPA 300  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYPA 300

1 MAKYSELEKGVIMQAFYWDVPSGGIWDITQKIPEDWDAGISAIWIPPSKMGCGAYS 60  
1 MAKYLEEGGLINQAFYWDVPMGGIWDITVAQKIPDWSAGISAIWIPPSKMGCGAYS 60  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
61 MGYDPDFFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVIADIVINHRAGGLEW 120  
121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
121 NPFVNDYTWTDFSKVSGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
181 QESYAALYLRISIGIDAWRFDYVKGYPAPVWVKDNLNMGWAGVGYWDTNVDVNLWYSSG 240  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYPA 300  
241 AKVDFPFLYKMDPEAFDNKNI PALVSALONGQTWVSERDPPKAVTFVANHTDIIWNKYPA 300

RESULT 8  
ABU03041  
ID ABU03041 standard; protein; 436 AA.  
XX  
AC ABU03041;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Alpha amylase polypeptide #2.  
XX  
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;

QY 301 YAFILTYEGOPTIFRYDYEBEWNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360  
 Db |||||  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEYTGNLGWDVXYVSSGWVLEAPADP 420  
 Db |||||  
 QY 421 ANGYGYSVMSYCGVG 436  
 Db |||||

RESULT 9  
 ABU03058  
 ID ABU03058 standard; protein; 436 AA.  
 AC  
 DT 21-JAN-2003 (first entry)  
 DE Alpha amylase polypeptide #19.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 OS Unidentified.  
 PN WO200268589-A2.  
 PD 06-SEP-2002.  
 PF 21-FEB-2002; 2002WO-US005068.  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Callen W, Richardson T, Frey G;  
 WPI: 2003-018556/01.  
 DR N-PSDB; ABX08429.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 PS Claim 30; Fig 16; 30lpp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch for modifying small molecules, for liquefying starch, for washing  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase

CC polypeptides of the invention  
 XX Sequence 436 AA;  
 SQ  
 Query Match 96.0%; Score 2360; DB 6; Length 436;  
 Best Local Similarity 95.6%; Pred. No. 1.1e-194;  
 Matches 417; Conservative 8; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVINQAFVWDPVSGGIWMDTIROKIPWYDAGISAIWIPASKMGSGAYS 60  
 Db |||||  
 QY 61 MGYPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYCMKVIADIVINHRAGGLEW 120  
 Db |||||  
 QY 121 NPFVNDYTWTFDSKVASKYKTANYLDPHPNELHAGDSGTGGYPDICHDKSWDQYMLWAS 180  
 Db |||||  
 QY 181 QESYAAVLRISIGIDAMRFYVKGYPWVVKDNLNMGWAVGEYWDVDAVLNWAYSSG 240  
 Db |||||  
 QY 241 AKVDFPALLYKMDAEPDNKNI PALVSALONGQTWVSRRDPKAVTFVANHTDIIWNKYPA 300  
 Db |||||  
 QY 301 YAFILTYEGOPTIFRYDYEBEWNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360  
 Db |||||  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEYTGNLGWDVXYVSSGWVLEAPADP 420  
 Db |||||  
 QY 421 ANGYGYSVMSYCGVG 436  
 Db |||||

RESULT 10  
 ABU03046  
 ID ABU03046 standard; protein; 436 AA.  
 AC ABU03046;  
 DT 21-JAN-2003 (first entry)  
 DE Alpha amylase polypeptide #7.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 OS Unidentified.  
 PN WO200268589-A2.  
 PD 06-SEP-2002.  
 PF 21-FEB-2002; 2002WO-US005068.  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Callen W, Richardson T, Frey G;

XX WPI; 2003-018656/01.  
 DR N-PSDB; ABX08417.  
 XX  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX  
 XX Claim 30; Fig 16; 301pp; English.  
 XX  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 XX Sequence 436 AA;  
 SQ  
 Query Match 95.9%; Score 2358; DB 6; Length 436;  
 Best Local Similarity 95.4%; Pred. No. 1.6e-194;  
 Matches 416; Conservative 9; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 DB 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 QY 61 MGVDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 DB 61 MGVDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 180  
 DB 121 NPFVNDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 180  
 QY 181 QESYAAYLRIGIDAMRFDYVKGYPVAVVVDLNMWGGWAGVGYWDTNVDALINWAYSSG 240  
 DB 181 QESYAAYLRIGIDAMRFDYVKGYPVAVVVDLNMWGGWAGVGYWDTNVDALINWAYSSD 240  
 QY 241 AKVDFPALLYKMBEAFPNKNIIPALVSALQNGQTVWSRDPKATFVANHDTDIWNKYPA 300  
 DB 241 AKVDFPALLYKMBEAFPNKNIIPALVSALQNGQTVWSRDPKATFVANHDTDIWNKYPA 300  
 QY 301 YAFLLTYEGOPTIFRYDYEELNKKDLKLIWIHENLAGGSTDIVYDNDDELIFVNGYG 360  
 DB 301 YAFLLTYEGOPTIFRYDYEELNKKDLKLIWIHENLAGGSTDIVYDNDDELIFVNGYG 360  
 QY 361 DKPLITYINLSSKAGRWVVPKFGACIHEYTGNLGGWVKYVSSGAVYLEAPAYDP 420  
 DB 361 SKPLITYINLSSKAGRWVVPKFGACIHEYTGNLGGWVKYVSSGAVYLEAPAYDP 420  
 QY 421 ANGOYGYSVNSYCGVG 436  
 DB 421 ANGOYGYSVNSYCGVG 436  
 RESULT 11  
 ABU03069  
 ID ABU03069 standard; protein; 436 AA.  
 XX  
 AC ABU03069;

XX 21-JAN-2003 (first entry)  
 DT  
 XX Alpha amylase polypeptide #30.  
 DE  
 XX Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 KW  
 XX Unidentified.  
 OS  
 XX WO200269589-A2.  
 PN  
 XX 06-SEP-2002.  
 PD  
 XX 21-FEB-2002; 2002WO-US005068.  
 PF  
 XX 21-FEB-2001; 2001US-0270495P.  
 PR  
 XX 21-FEB-2001; 2001US-0270496P.  
 PR  
 XX 14-MAY-2001; 2001US-0291122P.  
 XX  
 XX (DIVE-) DIVERSA CORP.  
 PA  
 XX Callen W, Richardson T, Frey G;  
 PI  
 XX WPI; 2003-018656/01.  
 DR  
 XX N-PSDB; ABX08440.  
 XX  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX  
 XX Claim 30; Fig 16; 301pp; English.  
 PS  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 XX Sequence 436 AA;  
 SQ  
 Query Match 95.9%; Score 2358; DB 6; Length 436;  
 Best Local Similarity 95.4%; Pred. No. 1.6e-194;  
 Matches 416; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 DB 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 QY 61 MGVDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 DB 61 MGVDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 180  
 DB 121 NPFVNDYDFFDLGEYDQKGTIVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 180  
 QY 181 QESYAAYLRIGIDAMRFDYVKGYPVAVVVDLNMWGGWAGVGYWDTNVDALINWAYSSG 240  
 DB 181 QESYAAYLRIGIDAMRFDYVKGYPVAVVVDLNMWGGWAGVGYWDTNVDALINWAYSSD 240  
 QY 241 AKVDFPALLYKMBEAFPNKNIIPALVSALQNGQTVWSRDPKATFVANHDTDIWNKYPA 300  
 DB 241 AKVDFPALLYKMBEAFPNKNIIPALVSALQNGQTVWSRDPKATFVANHDTDIWNKYPA 300  
 QY 301 YAFLLTYEGOPTIFRYDYEELNKKDLKLIWIHENLAGGSTDIVYDNDDELIFVNGYG 360  
 DB 301 YAFLLTYEGOPTIFRYDYEELNKKDLKLIWIHENLAGGSTDIVYDNDDELIFVNGYG 360  
 QY 361 DKPLITYINLSSKAGRWVVPKFGACIHEYTGNLGGWVKYVSSGAVYLEAPAYDP 420  
 DB 361 SKPLITYINLSSKAGRWVVPKFGACIHEYTGNLGGWVKYVSSGAVYLEAPAYDP 420  
 QY 421 ANGOYGYSVNSYCGVG 436  
 DB 421 ANGOYGYSVNSYCGVG 436  
 RESULT 11  
 ABU03069  
 ID ABU03069 standard; protein; 436 AA.  
 XX  
 AC ABU03069;



QY 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWNWGGWAVGEYDNTVDVNLWAYSSG 240  
DB 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWNWGGWAVGEYDNTVDVNLWAYSSD 240  
QY 241 AKVDFPFLYYKMDAEAFDNKNIIPALVSGALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
DB 241 AKVDFPFLYYKMDAEAFDNKNIIPALVSGALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
QY 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
DB 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
QY 361 DKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
DB 361 DKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
QY 421 ANGGYGSVWSYCGVG 436  
DB 421 ANGGYGSVWSYCGVG 436

RESULT 12  
ABU03054  
ID ABU03054 standard; protein; 436 AA.  
AC ABU03054;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Alpha amylase polypeptide #15.  
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
PN WO200268589-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 21-FEB-2002; 2002WO-US005068.  
XX  
PR 21-FEB-2001; 2001US-0270495P.  
XX  
PR 21-FEB-2001; 2001US-0270496P.  
XX  
PR 14-MAY-2001; 2001US-0291122P.  
XX  
PA (DIVE-) DIVERSA CORP.  
XX  
PI Callen W, Richardson T, Frey G;  
XX  
DR WPI; 2003-018656/01.  
XX  
DR N-PSDE; ABX08425.  
XX  
XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
XX liquefying starch, for textile desizing, for treating lignocellulosic  
XX fibers, and for producing high-maltose or high-glucose syrup.  
PS Claim 30; Fig 16; 301pp; English.  
XX  
CC The invention relates to a purified polypeptide with alpha-amylase  
CC activity and the polynucleotide encoding it. The polypeptide is useful  
CC for hydrolysing starch linkages, for catalysing the breakdown of a  
CC starch, for modifying small molecules, for liquefying starch, for washing  
CC an object, for textile desizing, for treating lignocellulosic fibers, for  
CC improving fibre properties, for enzymatic de-linking of recycled paper  
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging

CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
CC polypeptides of the invention  
XX  
SQ Sequence 436 AA;

Query Match 95.8%; Score 2356; DB 6; Length 436;  
Best Local Similarity 95.0%; Pred. No. 2.4e-194;  
Matches 414; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWMWDTIRKIPEDYDAGISATWIPASKMGMGAYS 60  
DB 1 MAKYSELEKGVIMQAFYWDVPSGGIWMWDTIRKIPEDYDAGISATWIPASKMGMGAYS 60  
QY 61 MGYPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGKVIADIVINHRAGGLEW 120  
DB 61 MGYPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGKVIADIVINHRAGGLEW 120  
QY 121 NPFVNDYTWDFSKVASGKYTANYLDPHNLHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
DB 121 NPFVNDYTWDFSKVASGKYTANYLDPHNLHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
QY 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWNWGGWAVGEYDNTVDVNLWAYSSG 240  
DB 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWNWGGWAVGEYDNTVDVNLWAYSSG 240  
QY 241 AKVDFPFLYYKMDAEAFDNKNIIPALVSGALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
DB 241 AKVDFPFLYYKMDAEAFDNKNIIPALVSGALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
QY 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
DB 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
QY 361 DKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
DB 361 DKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
QY 421 ANGGYGSVWSYCGVG 436  
DB 421 ANGGYGSVWSYCGVG 436

RESULT 13  
ABU03048  
ID ABU03048 standard; protein; 436 AA.  
XX  
AC ABU03048;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Alpha amylase polypeptide #9.  
XX  
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
PN WO200268589-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 21-FEB-2002; 2002WO-US005068.

XX 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 PA Callen W, Richardson T, Frey G;  
 DR WPI; 2003-018656/01.  
 DR N-PSDB; ABX08419.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 30; Fig 16; 301pp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX Sequence 436 AA;  
 SQ

Query Match 95.8%; Score 2356; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 2.4e-194;  
 Matches 415; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGLIMQAFYWDVPMGGIWDITROKIPEDWDAGISAIWIPASKMGGGYS 60  
 QY 61 MGYDPYDFDLGEYDQGTETREFGSKQELVNMINTAHAYGKVIADIVINHRAGDLEW 120  
 DB 61 MGYDPYDFDLGEYVQGTETREFGSKQELVNMINTAHAYGKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWDFSVASKYKATYLDHPHNLHAGSGTGGYPDICHKSDQYWLWAS 180  
 DB 121 NPFVGDYTWDFSVASKYKATYLDHPHNLHAGSGTGGYPDICHKSDQYWLWAS 180  
 QY 181 QESVAAVLRISIGIDAWDFYVKGYPAPVWVVDLNMWGGWAGVGYWDTNVDVILNWAYSSG 240  
 DB 181 QESVAAVLRISIGIDAWDFYVKGYPAPVWVVDLNMWGGWAGVGYWDTNVDVILNWAYSSG 240  
 QY 241 AKVDFPALYKWDFAFNKATIPALVSALONGQTVSRDPEKATVFNHNDTIIWKYPA 300  
 DB 241 AKVDFPALYKWDFAFNKATIPALVSALONGQTVSRDPEKATVFNHNDTIIWKYPA 300  
 QY 301 YAFILTYEGQPTIFRYDYEEWLNKDKLNIWIHNLHAGSGTIDIVYDNDLIFVRNGYG 360  
 DB 301 YAFILTYEGQPTIFRYDYEEWLNKDKLNIWIHNLHAGSGTIDIVYDNDLIFVRNGYG 360  
 QY 361 DKPGLITVINLGSAGRWVYVPKFACTHEVTGNLGNVYSSGWVILEAPAYDP 420  
 DB 361 SKPGLITVINLGSAGRWVYVPKFACTHEVTGNLGNVYSSGWVILEAPAYDP 420  
 QY 421 ANGQGYSVWSYCGVG 436  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60

Db 421 ANGQGYSVWSYCGVG 436

RESULT 14  
 ABU03066  
 ID ABU03066 standard; protein; 436 AA.  
 XX AC ABU03066;  
 XX DT 21-JAN-2003 (first entry)  
 XX DE Alpha amylase polypeptide #27.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX Unidentified.  
 OS WO200268589-A2.  
 PN 06-SEP-2002.  
 XX 21-FEB-2002; 2002WO-US005068.  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 PA (DIVE-) DIVERSA CORP.  
 PI Callen W, Richardson T, Frey G;  
 DR WPI; 2003-018656/01.  
 DR N-PSDB; ABX08437.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 30; Fig 16; 301pp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX Sequence 436 AA;  
 SQ

Query Match 95.6%; Score 2352; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 5.4e-194;  
 Matches 415; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60

Db 421 ANGQGYSVWSYCGVG 436

QY 61 MGYPDYDFDLGEYDQGTVEITRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 Db 61 MGYPDYDFDLGEYDQGTVEITRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
 Db 121 NPFVNDYTWDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
 QY 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
 Db 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
 QY 241 AKVPDFALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 Db 241 AKVPDFALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSGTDIVYDNDDELIFVRNGYG 360  
 Db 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSGTSIVYDSDENIFVRNGYG 360  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKFAAGCIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420  
 Db 361 SKPGLITYINLGSSKAGRWVYVPKFAAGCIHEYTGNLGWDVKWVDSSGWVYLEAPADHP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 Db 421 ANGQYGSVWSYCGVG 436

## RESULT 15

ABU03043  
 ID ABU03043 standard; protein; 436 AA.

AC ABU03043;  
 XX

DT 21-JAN-2003 (first entry)  
 DE Alpha amylase polypeptide #4.

KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.

OS Unidentified.

PN WO200268589-A2.

PD 06-SEP-2002.

PF 21-FEB-2002; 2002WO-US005068.

PR 21-FEB-2001; 2001US-0270495P.

PR 21-FEB-2001; 2001US-0270496P.

PR 14-MAY-2001; 2001US-0291122P.

PA (DIVE-) DIVERSA CORP.

PI Callen W, Richardson T, Frey G;  
 XX WPI; 2003-018656/01.

DR N-PSDB; ABX08414.

PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.

PS Claim 30; Fig 16; 301pp; English.

XX The invention relates to a purified polypeptide with alpha-amylase

CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper,  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 436 AA;

Query Match 95.6%; Score 2350; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 8e-194;  
 Matches 415; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIIMWDTIROKIPEDWDAGISATWIPPSKMGCGAYS 60  
 Db 1 MAKYSELEEGGLNQAFYWDVPMGGIIMWDTIAOKIPDWSAGISATWIPPSKMGCGAYS 60  
 QY 61 MGYPDYDFDLGEYDQGTVEITRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 Db 61 MGYPDYDFDLGEYDQGTVEITRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
 Db 121 NPFVNDYTWDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
 QY 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
 Db 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
 QY 241 AKVPDFALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 Db 241 AKVPDFALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSGTDIVYDNDDELIFVRNGYG 360  
 Db 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSGTDIVYDNDDELIFVRNGYG 360  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKFAAGCIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420  
 Db 361 SKPGLITYINLGSSKAGRWVYVPKFAAGCIHEYTGNLGWDVKWVDSSGWVYLEAPADHP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 Db 421 ANGQYGSVWSYCGVG 436

Search completed: June 29, 2004, 09:34:14  
 Job time : 61 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:32:02 ; Search time 21 Seconds

(without alignments)  
1997.119 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQGYGVMSYGVGV 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	492	20.0	492	2 AH2079	alpha-amylase (imp
2	487.5	19.8	494	2 AD0751	cytoplasmic alpha-
3	486.5	19.8	494	1 B45738	alpha-amylase (EC
4	475.5	19.3	549	1 A24549	alpha-amylase (EC
5	472.5	19.2	514	1 ALBSN	alpha-amylase (EC
6	466.5	19.0	549	1 A54541	alpha-amylase (EC
7	461.5	18.8	549	1 A24436	alpha-amylase (EC
8	459	18.7	548	1 ALBSF	alpha-amylase (EC
9	458.5	18.6	512	1 ALBSL	alpha-amylase (EC
10	454.5	18.5	518	1 A27705	alpha-amylase (EC
11	452.5	18.4	495	2 B90962	cytoplasmic alpha-
12	451.5	18.4	495	2 AD3038	alpha-amylase (EC
13	451.5	18.4	506	2 G98247	alpha-amylase (EC
14	447.5	18.2	495	1 A57338	alpha-amylase (EC
15	444.5	18.1	495	2 B5810	cytoplasmic alpha-
16	438	17.8	491	2 C56781	alpha-amylase (EC
17	420	17.1	493	2 S15713	alpha-amylase (EC
18	400	16.3	421	2 S10514	alpha-amylase (EC
19	398	16.2	423	2 T09942	alpha-amylase (EC
20	396.5	16.1	443	2 J01527	alpha-amylase (EC
21	390.5	15.9	445	2 S19990	alpha-amylase (EC
22	373	15.2	484	2 G95160	alpha-amylase (imp
23	372	15.1	484	2 F98026	alpha-amylase (EC
24	366.5	14.9	427	1 ALBHB	alpha-amylase (EC
25	365.5	14.9	438	1 ALBHB	alpha-amylase (EC
26	363.5	14.8	504	2 A55861	alpha-amylase (EC
27	357.5	14.5	438	2 S14937	alpha-amylase (EC
28	353	14.4	437	2 S14936	alpha-amylase (EC
29	351.5	14.3	429	1 J50406	alpha-amylase (EC

30 348.5 14.2 428 2 T05521 alpha-amylase (EC  
31 347.5 14.1 439 2 T02956 alpha-amylase (EC  
32 347.5 14.1 440 2 S14958 alpha-amylase (EC  
33 347 14.1 428 2 S10013 alpha-amylase (EC  
34 347 14.1 434 2 S12775 alpha-amylase (EC  
35 341.5 13.9 435 2 S12625 alpha-amylase (EC  
36 341.5 13.9 435 2 JC7137 alpha-amylase (EC  
37 340.5 13.8 437 2 S07040 alpha-amylase (EC  
38 339 13.8 826 2 E96720 probable alpha-amy  
39 336 13.7 437 2 J09446 alpha-amylase (EC  
40 336 13.7 437 2 J09446 alpha-amylase (EC  
41 335.5 13.6 413 1 ALMT3 alpha-amylase (EC  
42 294 12.0 412 2 C96789 protein T23E18.6  
43 288.5 11.7 547 2 A32803 glucan 1,4-alpha-m  
44 278 11.3 551 2 S05667 glucan 1,4-alpha-m  
45 253.5 10.3 710 2 S63598 cyclomaltodextrin

#### ALIGNMENTS

##### RESULT 1

AH2079

alpha-amylase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C:Accession: AH2079

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anai

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2079

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-492 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA73889.1; PID:gl7131281; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr2190

C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 20.0%; Score 492; DB 2; Length 492;

Best Local Similarity 27.8%; Pred. No. 4.6e-28;

Matches 147; Conservative 81; Mismatches 153; Indels 148; Gaps 28;

QY 8 EKGGVIMQAFYWDVPSGGIWDTIROKIPEDWDAGISAIWIPPSKMGCGAYSMGYDPTD 67

Db 3 QMGTMMQVPHYIPNDGNLWSKVEASAPELADAGTAMLPAYKGFAGSFDVGYGYVD 62

QY 68 FFDLGEYDQGTVEFRFGSKQELVNNINTAHAYGMKVIADIVINHAGDGL---EWNPF 123

Db 63 LFDLGEFDQKGSVRTKYGTQQYLDVAVSLQTHGLQYADAVLNHKGMDAVETPKATFF 122

QY 124 VND-----YTWDPSKVASGKYTANY-----LDHFHNEHLAGDSGT- 159

Db 123 PQDRLNPKGGLQDIKTYTHYFPFG-RQGY-SNFEWHWHHFDVAVD--NEYNSGDRSTV 178

QY 160 -----FGYYP-----DICHKXSDQ---YW-LWASQESVAALRSIGD 194

Db 179 YLLEGNFDDYVALEKGNFAYLMGCDLDFQNEWRGVEVYWGKWC-----LDTTKVD 230

QY 195 AMRFDYKGYAPVWVKDNLNMGWNA-----VGEYWDNTNVDVLNWL---AYSSGAKVFD 245

Db 231 GFRIDAIKHISTWFFPEWIDALERHAGKOLFVWGEYWNIDINTLL-WYWDVAVRGMKSVFD 289

QY 246 FALYYKMDA-----FDKNKIPALVSALQNGQTVVSRDPFKAVTFVAHNDTII----- 294

Db 290 VPLHYNFHQSAGSGGNYDMRRI-----LDG-TMMQQRPTAVTFVENHDSQPLAES 341

QY 295 ----WNKYPAYAFI-LTYEGQPTIFYRD-----YSEWLNKDK-----LKMLI 331

Db 342 VVEPWFKPLAYAILLRQEGYPCVFHADYGAIEDW-GKGNRYNIFMPGHRWIIDKLL 400  
Qy 332 WIHENLAGGSTDIVYDNDLI-FVRNGYDGPGLITVNLGSSKAGRWVYV---PKFA 386  
Db 401 YARKHAYG-PQNYLDHNTTIGTWLGDADHPQGMVIMDSGSGIKMVEGKPTKFI 459  
Qy 387 GACIHEYTONLGWDVYKYYSGWVLEAPAYDPANGQY-----GYSW 430  
Db 460 DLTEH-----IKEAVYTNWGW-----GPRCLGSGSVW 489

RESULT 2  
AD0751  
cytoplasmic alpha-amylase [imported] - Salmonella enterica subsp. enterica serovar Typhi  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 03-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AD0751  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AD0751  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-494 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD05711.1; PID:g16503204; GSPDB:GN00176  
C:Genetics:  
A:Gene: SY2171  
C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 19.8%; Score 487.5; DB 2; Length 494;  
Best Local Similarity 28.7%; Pred. No. 9.9e-28;  
Matches 139; Conservative 71; Mismatches 156; Indels 119; Gaps 21;  
Qy 9 KGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPASKMGAGYSGVDPYDF 68  
Db 2 KNPTLLQYFHWYDGGKLSLAERADGLNDIGINWMLPPACKGASGYSVGYDYL 61  
Qy 69 FDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKIADIVINHRAGDGLWNPV 124  
Db 62 FDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKIADIVINHRAGDGLWNPV 121  
Qy 125 -NDYT-----WTDFSKVA-SGKYTANYLDHF-----PNEHAGDSGTF- 160  
Db 122 QDRTQIDDDNIIECEGWTRYTFPPARAGQSNFIWDYHCFSGIDHIENPDE-----DGIFK 176  
Qy 161 -----GGYDICHDK-SWD-----QYWL-WASQESYAAYLSIG 192  
Db 177 IVNDYTGDNQDQVDDSLGNFDYLMGENIDFRNHAVENTEIKYWARVWMEQTHC----- 229  
Qy 193 IDAWRFYVKGYPWVVKMLNMGWGA-----VGEYWDNTVDVNLWAYSSGAK--VF 244  
Db 230 -DGRFLDAVKHHPANFYKEMIEHVAQVAPKPLFVAEYWSHEVDKLTQYIDQADGKTMLF 288  
Qy 245 DFALYKMDA-----FDKNKIPALVSALONGQTVSRDPFKAVTFVANHDTII----- 294  
Db 289 DAPLQMKFHEASRQAGYDMRHFT-----GTLVEADPFHATVILVANHDTQLOALE 340  
Qy 295 -----WNKYPAYAFILTYE-QOPTIFYED-----EMLNKD-----KLNLIW 332  
Db 341 APVEPWFKPLAYAILLRQEGYPCVFHADYGAIEDW-GKGNRYNIFMPGHRWIIDKLL 400  
Qy 333 IHENLAGGSTDIVYDNDLI-FVRNGYDGPGLITVNLGSSKAGRWVYVFPAGACIH 391  
Db 401 ARQRFAGH-IQTLFFDHPNCIAFSRSGTEENPGCVVLSNGDDGDKETLLGLDNYANKTW 459  
Qy 392 EYTGK 396  
Db 460 DFLGN 464

RESULT 3  
B45738  
alpha-amylase (EC 3.2.1.1), cytosolic - Salmonella typhimurium  
N:Alternate names: 1,4-alpha-D-glucan gluconohydrolase  
C:Species: Salmonella typhimurium  
C:Date: 07-Apr-1994 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
C:Accession: B45738  
R:Raha, M.; Kawagishi, I.; Mueller, V.; Kihara, M.; Macnab, R.M.  
J. Bacteriol. 174, 6644-6652, 1992  
A:Title: Escherichia coli produces a cytoplasmic alpha-amylase, amyA.  
A:Reference number: A45738; MUID:93015717; PMID:1400215  
A:Accession: B45738  
A:Molecule type: DNA  
A:Residues: 1-494 <RAH>  
A:Cross-references: GB:L01643; NID:g154043; PIDN:AAA27110.1; PID:g154045  
C:Genetics:  
A:Gene: amyA  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
A:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
C:Keywords: cytosol; glycosidase; hydrolase; polysaccharide degradation  
F:202-335/Domain: alpha-amylase core homology <AMY>  
F:239,265,332/Active site: His, Glu, Asp #status predicted

Query Match 19.8%; Score 486.5; DB 1; Length 494;  
Best Local Similarity 28.2%; Pred. No. 1.2e-27;  
Matches 140; Conservative 75; Mismatches 159; Indels 123; Gaps 21;  
Qy 9 KGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPASKMGAGYSGVDPYDF 68  
Db 2 KNPTLLQYFHWYDGGKLSLAERADGLNDIGINWMLPPACKGASGYSVGYDYL 61  
Qy 69 FDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKIADIVINHRAGDGLWNPV 124  
Db 62 FDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKIADIVINHRAGDGLWNPV 121  
Qy 125 -NDYT-----WTDFSKVA-SGKYTANYLDHF-----PNEHAGDSGTF- 160  
Db 122 QDRTQIDDDNIIECEGWTRYTFPPARAGQSNFIWDYHCFSGIDHIENPDE-----DGIFK 176  
Qy 161 -----GGYDICHDK-SWD-----QYWL-WASQESYAAYLSIG 192  
Db 177 IVNDYTGDNQDQVDDSLGNFDYLMGENIDFRNHAVENTEIKYWARVWMEQTHC----- 229  
Qy 193 IDAWRFYVKGYPWVVKMLNMGWGA-----VGEYWDNTVDVNLWAYSSGAK--VF 244  
Db 230 -DGRFLDAVKHHPANFYKEMIEHVAQVAPKPLFVAEYWSHEVDKLTQYIDQADGKTMLF 288  
Qy 245 DFALYKMDA-----FDKNKIPALVSALONGQTVSRDPFKAVTFVANHDTII----- 294  
Db 289 DAPLQMKFHEASRQAGYDMRHFT-----GTLVEADPFHATVILVANHDTQLOALE 340  
Qy 295 -----WNKYPAYAFILTYE-QOPTIFYED-----YEELNK-----DKLNLIW 332  
Db 341 APVEPWFKPLAYAILLRQEGYPCVFHADYGAIEDW-GKGNRYNIFMPGHRWIIDKLL 400  
Qy 333 IHENLAGGSTDIVYDNDLI-FVRNGYDGPGLITVNLGSSKAGRWVYVFPAGACIH 391  
Db 401 ARQRFAGH-IQTLFFDHPNCIAFSRSGTEENPGCVVLSNGDDGDKETLLGLDNYANKTW 459  
Qy 392 EYTGK 408  
Db 460 DFLGN 472

RESULT 4  
A24549  
alpha-amylase (EC 3.2.1.1) precursor - Bacillus stearothermophilus (strain NZ-3)  
N:Alternate names: 1,4-alpha-D-glucan gluconohydrolase  
C:Species: Bacillus stearothermophilus



Wed Jun 30 10:27:03 2004

C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: B85810  
 R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: B85810  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-495 <STO>  
 A:Cross-references: GB:AE005174; NID:gl2516028; PIDN:AAGS6942.1; GSPDB:GN00145; UWGP:Z30  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: amyA  
 C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 18.1%; Score 444.5; DB 2; Length 495;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-24;  
 Matches 131; Conservative 67; Mismatches 140; Indels 119; Gaps 21;

QY 13 IMCAFVWDVPSGIGIMWDTIRQIPKIPWYDAGISAIWIPKSGMGAYSGYDYPYDFDLG 72  
 DB 6 LLQCFHWYYPGGKLPWELAEADGFNDIGINWMLPPAYKASGSGYSGYDSDYDLFDLG 65  
 QY 73 EYDQGTVTRETSKQELNMINTAHYGMKVIADIVINHRAGDLE-----WNEFFVNDY 127  
 DB 66 EFDQKGSIPKTYGDKVQLLAADALKRNDIAVLDDVNNHKGADKEAIRVQRVNADDR 125  
 QY 128 T-----WTDFSKVA-SGKYTANYLDF-----HPNELHAGDSGTF----- 160  
 DB 126 QTDEEIIIECEGWRITFPFARAGQSQFIWDFKCFSGIDHIENPNE-----DGIFKIYND 180  
 QY 161 ---GGYPDICHDK--SWD-----QYWL-WASQESYAAYLRSIGIDAM 196  
 DB 181 YTCEGRNDQVDDDELGNFDYLMGENIDFRNHAVTEEEKYWARWMEQTQC-----DGF 232  
 QY 197 RFDYVKGYPWVKDMLNHWGGWA-----VGEYMDTNVDVNLWAYSSGAK--VDFDFAL 248  
 DB 233 RLDAVKHPIPAWFYKEWIEHVQEVAPKPLFIVABYSHEVDKLTQYIDQVEGKTNLFDAPL 292  
 QY 249 YIKWDEA-----FONKNIPALVSLQNGQTVVSRDPFKAVTFVANHDTDI----- 294  
 DB 293 QMKFHEASRGRNYD-----MTQIFTG-TLVEADXFHAVIXVANHDTQXLALEAPVE 344  
 QY 295 -WNKYPAYAFILTYE-GQPTIFYRD-----YEELNK-----DKLNLIWIHEN 336  
 DB 345 PWFKPLAYALLIRENGVPSVFPDLYGAHYEDVGGDGTYPIDMPIIEQLDELILARQ 404  
 QY 337 LAGSSTDIVYNDDEL-FVRNGYGDKPLITYINLG 372  
 DB 405 FARG-VOTLFFHPNCIAFSRGTDEVPCCGVVMSNG 440

Search completed: June 29, 2004, 09:36:16  
 Job time : 23 secs



Db 295 SFHDASKQGGDFDMRSIFDGLSLVAV-----PDHAVTLVDNHDTPQLQSLAPVE 344  
Qy 295 -WNKYPAYAFI-LTYEGQPTIFRD-----YEELWLNK-----DKLKNLIWIHEN 336  
Db 345 PWFKPLAYAILLREBEGVPCVFDLPFGTSTDTGNDGNEYKIDIPAIACLPLKLEARSR 404  
Qy 337 LAGGSTDIVYDNDDELIFVRNGYGDKEPLITYINLGSSKAGRWVYVPKAGACIHEYTCN 396  
Db 405 FANGPQTDIFDDASCIARHGTADAPGCVVMSNGEPGEKQADLGPBRAGSVWRDFLGH 464  
RESULT 13  
G98247  
cytoplasmic alpha-amylase (1,4-alpha-D-glucan glucanohydrolase) [imported] - Agrobacterium  
C;Species: Agrobacterium tumefaciens  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
C;Accession: G98247  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: G98247  
A;Status: preliminary  
A;Molecule type: DNA  
A;Gene: AGR\_L1863  
A;Residues: 1-506 <KUR>  
A;Cross-references: GB:AE007870; PIDN:AAK89505.1; PID:g15159379; GSPDB:GN00170  
C;Genetics:  
A;Map position: linear chromosome  
C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
Query Match 18.4%; Score 451.5; DB 2; Length 506;  
Best Local Similarity 27.3%; Pred. No. 4.1e-25;  
Matches 131; Conservative 70; Mismatches 168; Indels 111; Gaps 15;  
Qy 10 GGVMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGKGAYSMGYDYPDF 69  
Db 14 GRLLQFFHWYDYGKLSVAEAKESLAKGIGTDVWLPAYKGAAGSYGYDITDLF 73  
Qy 70 DLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVNHRAGGDLWNPFFV---- 124  
Db 74 DLGEYDQKGTATKYGDRAALSHAGKTLKDGIRVHDVVLNHNKGADEKVRVRVNP 133  
Qy 125 ---NDYTWTFDSKVASGKT-----ANVL 145  
Db 134 DRTDIDDEDFPALATRTFFGRNGHSHKFTWDLKCFSGVDHIEPTEDGIFRLWNEYG 193  
Qy 146 DFPNELHAGDSGTF----GGYDICHDKSDQ--YW-LWASQESYAAAYLRSIGIDAWRF 198  
Db 194 DGEWNEEVDOENGDFDLMGADVEFENRAVYELKYGWRWLSQ-----VQVDGFL 245  
Qy 199 DIVKGYAPVWVKDLNWWCG-----WAGVETWDTNVDVNLW--AYSSGAKVDFALYY 250  
Db 246 DAAKHIPAWFFRDVGHMRETVDPLFVVAEYWHPDLEALKSYLBVLDKQLMLFDVALHH 305  
Qy 251 KMDEA-----FDKNIPALVQNGQTVVSRDPPKAVTFVANHTDII----- 294  
Db 306 SFHDASKQGGDFDMRSIFDGLSLVAV-----PDHAVTLVDNHDTPQLQSLAPVE 355  
Qy 295 -WNKYPAYAFI-LTYEGQPTIFRD-----YEELWLNK-----DKLKNLIWIHEN 336  
Db 356 PWFKPLAYAILLREBEGVPCVFDLPFGTSTDTGNDGNEYKIDIPAIACLPLKLEARSR 415  
Qy 337 LAGGSTDIVYDNDDELIFVRNGYGDKEPLITYINLGSSKAGRWVYVPKAGACIHEYTCN 396  
Db 416 FANGPQTDIFDDASCIARHGTADAPGCVVMSNGEPGEKQADLGPBRAGSVWRDFLGH 475  
RESULT 14  
A45738  
alpha-amylase (EC 3.2.1.1), cytosolic - Escherichia coli (strain K-12)  
N;Alternate names: 1,4-alpha-D-glucan glucanohydrolase

C;Species: Escherichia coli  
C;Date: 07-Apr-1994 #sequence\_revision 31-Oct-1997 #text\_change 01-Mar-2002  
C;Accession: D64956; A45738  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: D64956  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-495 <BLAT>  
A;Cross-references: GB:AE000285; GB:U00096; NID:g1788229; PIDN:AAK74994.1; PID:g1788236;  
A;Experimental source: strain K-12, substrain M61655  
R;Raka, M.; Kawagishi, I.; Mueller, V.; Kihara, M.; Macnab, R.M.  
J. Bacteriol. 174, 6644-6652, 1992  
A;Title: Escherichia coli produces a cytoplasmic alpha-amylase, amyA.  
A;Reference number: A45738; MUID:93015717; PMID:1400215  
A;Accession: A45738  
A;Molecule type: DNA  
A;Residues: 1-18,'SS',21-108,'V',110-148,'E',150-233,'I',235-495 <RAH>  
A;Cross-references: GB:L01642; NID:g146021; PIDN:AAA23810.1; PID:g146023  
C;Genetics:  
A;Gene: amyA  
C;Function:  
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A;Pathway: glycogen/starch degradation  
C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
C;Keywords: cytosol; glycosidase; hydrolase; polysaccharide degradation  
F;202-335/domain: alpha-amylase core homology <AMY>  
F;104,204,239/Binding site: calcium (Asn, Asn, His) #status predicted  
F;235,265,332/Active site: Asp, Glu, Asp #status predicted  
Query Match 18.2%; Score 447.5; DB 1; Length 495;  
Best Local Similarity 28.7%; Pred. No. 7.8e-25;  
Matches 131; Conservative 68; Mismatches 139; Indels 119; Gaps 21;  
Qy 13 IMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGKGAYSMGYDYPDFDLG 72  
Db 6 LLQCFHWYDYGKLSVAEAKESLAKGIGTDVWLPAYKGAAGSYGYDITDLF 65  
Qy 73 BYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVNHRAGGDLWNPFFV---- 127  
Db 66 EFDQKGSIPKYGDRAALSHAGKTLKDGIRVHDVVLNHNKGADEKVRVRVNP 125  
Qy 128 T-----WTFDSKVA-SGKTYANYLDF-----HPNELHAGDSGTF----- 160  
Db 126 TQDDEIEIECEGWTRYTFPARAGYSQFIWDFKFCSDIHENPDE-----DGIFKIVND 180  
Qy 161 ---GGYDICHDK-SWD-----OYWL-WASQESYAAAYLRSIGIDAW 196  
Db 181 YTEGMDNDQVDDBLGNFDYLMGENIDFRNHAVTEIKYWARWMEQTQC-----DGF 232  
Qy 197 RFDYKGYAPVWVKDLNWWCGA-----VGEYWDTNVDVNLWYSSGAK--VFDFAL 248  
Db 233 RLDAVKHIIPAWFKIEWIEHVEVAPKPLFIVAYNSHEVDKLTQYIDQVEGKTMFLDAPL 292  
Qy 249 YKKDEA-----FDKNIPALVQNGQTVVSRDPPKAVTFVANHTDII----- 294  
Db 293 QMKFHEASRMGRDYD-----MTQIFTG-TLVEADPFHATILVANHTQPLQALEAPVE 344  
Qy 295 -WNKYPAYAFI-LTYEGQPTIFRD-----YEELWLNK-----DKLKNLIWIHEN 336  
Db 345 PWFKPLAYAILLREBEGVPCVFDLPFGTSTDTGNDGNEYKIDIPAIACLPLKLEARSR 404  
Qy 337 LAGGSTDIVYDNDDELIFVRNGYGDKEPLITYINLGSSKAGRWVYVPKAGACIHEYTCN 372  
Db 405 FANGPQTDIFDDASCIARHGTADAPGCVVMSNGEPGEKQADLGPBRAGSVWRDFLGH 440  
RESULT 15  
B65810  
cytoplasmic alpha-amylase [imported] - Escherichia coli (strain O157:H7, substrain EDL93;

11	QY	GVINQAFVWDVPSGGIWWDTTRQIKIPEDWDAGISAIWIIPASKXGNGGAYSMGVDPYDFFD	70
40	Db	GTMMQYFEWYLPDNGNHNELNSDASNLKSGITAVWIPANKG-ASQNDVGAYGAYDLYD	98
71	QY	LCEYDOKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGD-----LEWNP-	122
99	Db	LGEFNQKGTETRYKGTSQLQAAVTSLKNGIGIYQVGDVWNHKGKADATEMVRAVEVNFN	158
123	QY	-----FVNDYT-----WTDFSKVASGK-----YTANYLDPHNELHAGDSGTFGGVPD	165
159	Db	NRNQEVTTGEYITAEATRFDFPGRGNTHSSPKFWRMYHFDGVDWDQSRLLNNRIYKFRG--	215
166	QY	1CHDKSWD-----QYWLWAS-----QESYAALRSIGIDAMRZDVKYG	203
216	Db	-----HGKAWDEYDTENGNYDILAYADIMDHPVEVVELRWGVWYTWTLGLDGRIDAVKH	273
204	QY	YAPVWVKDWLNNWGG-----WAVEYWDTNVDV-----LNWYSSGAKYDFDFALYK	251
274	Db	IKYSTRDWINHVRSATGSKMFEAWEFQRNDLGAENIYKQTNWHS-----VFDVPLHYN	329
252	QY	MDEA-----FDKNKIIPALVSALQNGOTVYSRDPFKAVTFVANHDTD-----IWN	296
330	Db	LYNASKSGNYDMENI-----PNG-TTVQRHPSHAVTFVDNHSQDPEALESFVEWF	381
297	QY	KYPAYAFILTYE-QOPTIFRDRVEW-----LNKOKLXNLIWIHNLAGGSTDIVVDN	349
382	Db	KPLAYALTLTREQYPSVFYGDYGIPTGCVPMRSGKIDPILBAROKYAYGQN-DYLDH	440
350	QY	DEL1-FVRNGYCDXP-----GLTYINLGSSKAGRWVVPK-PAGACTHEYTNLGGWWDVYV	405
441	Db	HNIICWTREGNHTAHNPSGLATIMSDGAG-CSKMFVGRNKGAGQVMSDITGNETGIVT--	497
406	QY	YSSGW	410
498	Db	NADGW	502

RESULT 11  
 B90962  
 cytoplasmic alpha-amylase [imported] - Escherichia coli (strain O157:H7, substrain RMD  
 C-species: Escherichia coli  
 C-date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C-accession: B90962  
 R-Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.;  
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A-title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
 A-reference number: A99629; PMID:21156231; PMID:11258796

A.Accession: B30962  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-495 <HAY>  
A.Cross-references: BIDN:BA83089.1; PID:gi3362134; GSPDB:GN00154  
A.Experimental source: strain O157:H7, substrain RMD 0509952  
C.Genetics:  
C.AGene: EC8266  
C.Superfamily: alpha-amylose, amylioliquefaciens type; alpha-amylose core homolo  
Query Match 18.4%; Score 452.5; DB 2; Length 495;  
Best Local Similarity 28.9%; Pred. NO. 3.4e-25;  
Matches 132; Conservative Matches 139; Mismatches 139; Indels 119; Gaps 21;

```

161 QY ---GGYPDICHDK---SWD-----QYWL-WASQESYAALRSIGIDAW 196
181 Db YTGEGWNDQVDELGNFDYLMNGENIDFRNHAVTEETIKYWARVWMEQTC-----DGF 232
197 QY RFDYVKGVAWVKWDLNWWGWA-----VGEYWDNTVDVLNWAYSSGAK--VPDFAL 248
233 Db RLDVAKGHPAFWFKYEWIEHVQEVAPKPLFLIVABYWSHEVDKLTQVIDQVEGKTMFLDAPL 292
249 QY YYKWDSEA-----PDKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDII----- 294
293 Db QMKFHEASRMGRNYD-----MTQIFTG-TLVEADPFHAVTLVANHDTQPLQALEAPVE 344
295 QY -WNKYPAVAFILTVE-GOPTIFYRD-----YSEWLNK-----DKLNLIWIHEN 336
345 Db PWFKPLAYALILLRENGVSVFFPDLYGAHYEDVGGDGQGYPIDMPIEQDELILARQR 404
337 QY LAGGSTDIIVYDNDELI-FVRNGYGDKPGLIYYINLG 372
405 Db FAHG-VQTLFFDHPNCIAFSRSGTDEYPCGVVMSNG 440

RESULT 12
AD3038
alpha-amyase amyA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AD3038
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD3038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL44722.1; PID:gi7742354; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: amyA
A:Map position: linear chromosome
C:Superfamily: alpha-amyase, amyloliquefaciens type; alpha-amyase core homolo

```

Query Match	18.4%;	Score 451.5;	DB 2;	Length 495;
Best Local Similarity	27.3%;	Pred. No. 4e-25;		
Matches 131;	Conservative 70;	Mismatches 168;	Indels 111;	Gaps 15;
QY	10	GGVIMQAFYVDVPSGGIWDTTTQRIKPEWYDAGISAIWIPSPASKMGQAYSGMGVDPDFF	69	
DBb	3	GRTLQPFHYYPDGGKLMSEVAEKAESLAKGITDVMPLPAYKGAAGYSGVGYDYLDF	62	
QY	70	DLGEYDQKGVTEPRFGSKBELVMNINTAHAYGMKYIADIVINHAGGLEWNPV	124	
DBb	63	DLGEFDQKGVTAIKYGDRALAEHAGKTLKDNIGIRIVHDVNLNKMGADEKEKVRVRWP	122	
QY	125	---NDYTWTDPSKVASGKVT-----	---ANYL	145
DBb	123	DDRTDIDDEDPALAYTRTFPRNGKHSKFTWDLKCFSGVDHIEEPTDGI FELVNEYG	182	
QY	146	DFHPNELHAGDSGTG---GGVPDICHDKSWDQ--YW-LWASQESYAAALRSIGIDAWRF	198	
DBb	183	DGEWNEEVQENGDFYLMGADVFEFNRAVYBELKYWGRWLSEQ-----	---VQVDG	234
QY	199	DYVKGYAPVWVKDNLNMGWG-----WAVEGYTWDTNVDVNLN---AYSSGAKVDFPLYY	250	
DBb	235	DAAKHIFAPFFRDVWGHMRETDPDLFVVAEYWHFDLSALKSYLIELVDKQMLJFDVALHH	294	
QY	251	KMDEA-----FDNKNI--PALVSLQNGQTVVSRDPFKAVTFVANHDTDII-----	---294	

### Query Match

```

Db 40 GTMWOYFEWLPDDGTLTKVANEANNLSSUGITATLWLPAYKGTSS--DVGYGVYDLYD 98
QY 71 LGEYDQKGTETFRFGSKOELNMINTAHAYGMKVIADIVINHRAGGD--LEW-----NPF 123
Db 99 LGEPNKGAVRTKYGTAKYQLQAIAHAAGQVYADVDFHKGADGTEWDAVEVNP 158
QY 124 VND-----YTWDFSKVASGYKTYANYLDFHFNELHAG-----155
Db 159 DRNOBISGTYQIAWTKFDFPGRGN---TYSFVKRWYTHFGVDWDSRKLRIYKPRGI 215
QY 156 -----DSGTG-----GYDICH--KSDQYVWLMASQESYAYLRSIGIDA 195
Db 216 GKADWEVDVTENGNDYLMYADLDWDHDEVTTELKSWGK-----YVNTNIDG 264
QY 196 WRFDYVKGAPVWVKDLNWMG-----WAVEGWDTNVDVAVLNWYSSGA--KVPDFA 247
Db 265 FRLDVAGTKHFKFPPDLSDVRSQTKPLFTVGEYWSYDINKLHNYIMKTNGTMSLFDAP 324
QY 248 LYKMDER-----FDKNKIPALVSALQNGQTVVSERDPFKAVTFVANHDTD-----292
Db 325 LHNKYTTASKSGTDFMRLMT-----NLMKQDQTLAVTFVNDHDTFPGQALQSWV 376
QY 293 IINWKPAYAFILT-YEGOPTIFYRDYE--EWLN-----KDKLNLIIWHENLAGSSTIV 345
Db 377 DPWFKPLAVAFILTRQEGYPCVFGYGYGIPQYNIPSLKSKIDPLIIARRDYAYG--TQHD 435
QY 346 YVDNDELI-FVRNGYDGP--GLITYINLGSSKAGRWYVVK--FAGACIHEVTGNLGGWV 401
Db 436 YLDHSDIIGWTREGVTEKPGSLAALITDGP--GSKWMTVGKQHAGKVFYDLTGNRSTV 494
QY 402 DKVYVSSGWYLEAPAYDPANGQY-----GYSVW 430
Db 495 T--INSDG-----GEFKVNGSGSVW 514

```

## RESULT 8

## ALBSL

```

alpha-amylase (EC 3.2.1.1) precursor - Bacillus stearothermophilus (strain DY-5) plasmid
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase
C:Species: Bacillus stearothermophilus
C:Date: 30-Jun-1987 #sequence revision 30-Jun-1987 #text_change 16-Feb-1997
C:Accession: A91999; B91999; A91804; A00845
R:Ihara, H.; Sasaki, T.; Tsuboi, A.; Yamagata, H.; Tsukagoshi, N.; Uda, S.
J. Biochem. 98, 95-103, 1985
A:Title: Complete nucleotide sequence of a thermophilic alpha-amylase gene: homology bet
A:Reference number: A91999; MUID:86008166; PMID:3876333
A:Accession: A91999
A:Molecule type: DNA
A:Residues: 1-548 <IH1>
A:Cross-references: GB:X02769
A:Experimental source: plasmid pH1300 from strain DY-5
A:Accession: B91999
A:Molecule type: protein
A:Residues: 35-48 <IH2>
A:Experimental source: strain DY-5
R:Tsukagoshi, N.; Iritani, S.; Sasaki, T.; Takemura, T.; Ihara, H.; Ido, Y.; Yamagata,
J. Bacteriol. 164, 1182-1187, 1985
A:Title: Efficient synthesis and secretion of a thermophilic alpha-amylase by protein-pr
A:Reference number: A91804; MUID:86059211; PMID:2999073
A:Contents: pBAM101
A:Accession: A91804
A:Molecule type: DNA
A:Residues: 1-29,'Q',31-75,'W',77-122 <TSU>
C:Comment: Alpha-amylase genes have been found on plasmids and in multiple copies on the
C:Genetics:
A:Genome: plasmid
A:Start codon: GIG
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: alpha-amylase, amylioliquefaciens type; alpha-amylase core homology
C:Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac

```

```

F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-548/Product: alpha-amylase #status experimental <MAT>
F:335-368/Domain: alpha-amylase core homology <AMY>
F:139,237,272/Binding site: calcium (Asp, Asp, His) #status predicted
F:268,298,365/Active site: Asp, Glu, Asp #status predicted

```

Query Match 18.7%; Score 459; DB 1; Length 548;

Best Local Similarity 28.6%; Pred. No. 1.3e-25;

Matches 148; Conservative 67; Mismatches 162; Indels 140; Gaps 26;

QY 11 GVTMQAFYVDVPSGGIWNVTIRQKIPEDYDAGISAIWIPASKGMGAGYSMGYDYPDFD 70

Db 40 GTMWOYFEWLPDDGTLTKVANEANNLSSUGITATLWLPAYKGTSS--DVGYGVYDLYD 98

QY 71 LGEYDQKGTETFRFGSKOELNMINTAHAYGMKVIADIVINHRAGGD--LEW-----NPF 123

Db 99 LGEPNKGAVRTKYGTAKYQLQAIAHAAGQVYADVDFHKGADGTEWDAVEVNP- 157

QY 124 VNDYTWDFSKVASGYKTYANYLDFHFNELHAGSGTFFGYPDICHDKSWDQY-----WL 177

Db 158 -----SDRNOEISGTY-----QIAWTKFDFPGRGNTYSSFKRWYHFDGVDWD 201

QY 178 WASQESYAVALRSIGIDAM-----RFDY-----VKGYPWVVK-- 210

Db 202 ESKLSRIYKFRIG--KANDWEVDVTENGNDYLMYADLDWDHDEVTTELKSWGKVVNTT 260

QY 211 -----DNLNWMG-----WAVEGWDTNVDVAVLNWYSSGA--KV 243

Db 261 NIDGFRDLGLKHFKFPPDLSDVRSQTKPLFTVGEYWSYDINKLHNYIMKTNGTMSL 320

QY 244 FDALYYKM-----DEAPDNKNIIPALVSALQNGQTVVSERDPFKAVTFVANHDTDI 294

Db 321 FDAPLHNKFTASKSGGAFDMRLMT-----NLMKQDQTLAVTFVNDHDTNPAKRC 372

QY 295 -----WNKYPAYAFILT-YEGOPTIFYRDYE--EWLN-----KDKLNLIIWHENLAGSST 342

Db 373 SHGPRWFKPLAVAFILTRQEGYPCVFGYGYGIPQYNIPSLKSKIDPLIIARRDYAYG-T 431

QY 343 DIVYDNDDELI-FVRNGYDGP--GLITYINLGSSKAGRWYVVK--FAGACIHEVTGNLG 398

Db 432 QHDYLDHSDIIGWTREGVTEKPGSLAALITDAGRS--KMWYVGKQHAGKVFYDLTGNRS 490

QY 399 GWVDKYVYSSGWYLEAPAYDPANGQY-----GYSVW 430

Db 491 DTVT--INSDG-----GEFKVNGSGSVW 513

## RESULT 9

## ALBSL

alpha-amylase (EC 3.2.1.1) precursor [validated] - Bacillus licheniformis

N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase

C:Species: Bacillus licheniformis

C:Date: 30-Jun-1987 #sequence revision 24-Apr-1998 #text\_change 15-Sep-2000

C:Accession: A91997; B24549; A91796; A21663; I39774; I39772; A26151; S53788; A00844

R:Yuuki, T.; Nomura, T.; Tezuka, H.; Tsuboi, A.; Yamagata, H.; Tsukagoshi, N.; Uda, S.

J. Biochem. 98, 1147-1156, 1985

A:Title: Complete nucleotide sequence of a gene coding for heat- and pH-stable alpha-amyl

ases deduced from the DNA sequences.

A:Reference number: A91997; MUID:86111694; PMID:2418011

A:Accession: A91997

A:Molecule type: DNA

A:Residues: 1-162,'R',164-512 <YUU>

A:Cross-references: GB:X03236; NID:G39551; PIDN:CAA26981.1; PID:G39552

A:Experimental source: ATCC 27811

R:Gray, G.L.; Mainzer, S.E.; Ray, M.W.; Lamsa, M.H.; Kindle, K.L.; Carmona, C.; Requa, J.

J. Bacteriol. 166, 635-643, 1986

A:Title: Structural genes encoding the thermophilic alpha-amylases of Bacillus stearother

A:Reference number: A91817; MUID:86195857; PMID:3009417

A:Accession: B24549

A:Molecule type: DNA

A:Residues: 1-338,'G',340-348,'S',350-512 <GRA>

A:Cross-references: GB:M13256; NID:G142510; PIDN:AAA22240.1; PID:G142511

A:Experimental source: NCIB 8061

QY 123 -----FVNDY---TWTFDSKVASKYKANYLDHPHNEH----- 153  
Db 153 NRQETSEEIQAIAWIDFPFGRGN---TYSDFKWHYHFDGADWDESKISRIKFRGE 209  
QY 154 -----AGDSGTFG---GYPDICHD-----KSWDQYWLWASQESYAAVLRISIGIDA 195  
Db 210 GKAWDEWVSSENGNYDYLWADVDYDHPDVVAETKKWG---IW-----YANLSLDG 258  
QY 196 WRFDYKGYAPVWVKDWL-----NWWGWAAGEYWDNVDVAVLW---AYSGAKVFPFA 247  
Db 259 FRIDAARKHIFSLRDVQVQVQATCKEMFTVAEYQNNAGLENYLNKTSFNQSVFVDP 318  
QY 248 LYYKMDA-----FNNKNIIPALVSALQNGQTVVSRDPFKAVTFVANHDT----- 292  
Db 319 LHENLQAASSGGGYDMR-----LLDG-TVSHPEKAVTFVENHDTQCGSLESTV 370  
QY 293 IWNKYPAYAFILTYE-GOPTIFRYDY-----EWNLNKDKLNLIWIHENLAGGST 342  
Db 371 QTWFKLAYAFILTYRESGYPQVFGMYGKTGTSPEIPSLKDNIEPIKARKEYAYGPQ 430  
QY 343 DIVVYDNDLI-FVRNGYGD--KPGILITYINLGSSKAGWVVPKFAACIHEYTCNLGG 399  
Db 431 H-DYIDHPDVIGWTRGDSAAKSGLAALITDQPGSKMYAGLNAGETWIDITGNRSD 489  
QY 400 WDKYVYSSGW 410  
Db 490 TVK--IGSDGW 498

RESULT 6  
A:alpha-amylose (EC 3.2.1.1) precursor - Bacillus stearothermophilus (strain DNI1792)  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Bacillus stearothermophilus  
C:Date: 28-Oct-1994 #sequence\_revision 18-Aug-1995 #text\_change 13-Jun-1997  
C:Accession: A54541  
R:Jorgensen, P.L.; Poulsen, G.B.; Diderichsen, B.  
FEMS Microbiol. Lett. 77, 271-276, 1991  
A:Title: Cloning of a chromosomal alpha-amylose gene from Bacillus stearothermophilus.  
A:Reference number: A54541  
A:Accession: A54541  
A:Molecule type: DNA  
A:Residues: 1-549 <OR>  
A:Cross-references: GB:X59476  
A:Experimental source: chromosomal DNA of strain DNI1792  
C:Comment: Alpha-amylose genes have been found on plasmids and in multiple copies on the  
C:Genetics:  
A:Start codon: GTG  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Superfamily: alpha-amylose, amyloliquefaciens type; alpha-amylose core homology  
C:Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac  
F:1-34/Domain: signal sequence #status predicted <Sig>  
F:35-549/Product: alpha-amylose #status predicted <MAT>  
F:235-368/Domain: alpha-amylose core homology <AMY>  
F:139, 237, 272/Binding site: calcium (Asp, Asp, His) #status predicted  
F:268, 298, 365/Active site: Asp, Glu, Asp #status predicted

Query Match 19.0%; Score 466.5; DB 1; Length 549;  
Best Local Similarity 28.4%; Pred. No. 3.7e-26;  
Matches 146; Conservative 67; Mismatches 168; Indels 133; Gaps 25;

QY 11 GVIMQAFYDVPSSGIWDTTQKIPWYDAGISAIWIPPPASKMGAGYSGYDYPDFD 70  
Db 40 GMMQYFVWYLPDDGTLTKVANEANNLSIGITALWLPYKGTSTRS-DVGYGVYDLYD 98  
QY 71 LGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHAGGD-LEW-----NPF 123  
Db 99 LGEFNQKGVTRTKYKTAQYLQIAAHAGQVYADVVFHKGADGTENVDAVENPS 158  
QY 124 VND-----YTWTFDSKVASKYKANYLDHPHNEHAG----- 155

Db 159 DRNOEISGTVOIQAWTKDFPGRGN---TYSSPKRWYHFDGVDWDESKLSRIYKFRGI 215  
QY 156 -----DSGTFG-----GYPDICHD-KSWDQYWLWASQESYAAVLRISIGIDA 195  
Db 216 GKAWDEWVDIENGNYDYLWADLMDHPVVVTELKWKW-----YVNTINIDG 264  
QY 196 WRFDYKGYAPVWVKDWLWGG-----WAVEYWDNVDVAVLWYSSGA--KVDFDA 247  
Db 265 FRIDAARKHIFSPFPDWSYVRSQTGKPLFTVGEWYSYDINKLHNYITKTGDTMSLPDAP 324  
QY 248 LYYKMDA-----DEAFDNKNIIPALVSALQNGQTVVSRDPFKAVTFVANHDT----- 292  
Db 325 LHNFYFASKSGGAFNRLMT-----NTLMKQDPTLAVTFVDNHDTEPQALOSWV 376  
QY 293 IWNKYPAYAFILTYE-GOPTIFRYDYE--EWNLNKDKLNLIWIHENLAGGSTDI 345  
Db 377 DPWFKPLAYAFILTYRQEGYPCVFGYDYGIPQYNIPSLKSKIDPILLIARRDYAVG-TQHD 435  
QY 346 YVDNDELII-FVRNGYGDK--GLITYINLGSSKAGWVVPKFAACIHEYTCNLGGW 401  
Db 436 YLHSDIIGWTRGEGTEKPSGLAALITDQPG-GSKMYVGVKQHAGKVFDYDLTGNRSDTV 494  
QY 402 DKYVYSSGWYLEAPAYDPANGQY-----GYSVW 430  
Db 495 T--INSDGW-----GBFKVNGSGSVW 514

RESULT 7  
A:alpha-amylose (EC 3.2.1.1) precursor - Bacillus stearothermophilus plasmid pAT5  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Bacillus stearothermophilus  
C:Date: 05-Jun-1987 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
C:Accession: A24436; I39777  
R:Nakajima, R.; Imanaka, T.; Aiba, S.  
J. Bacteriol. 163, 401-406, 1985  
A:Reference number: A24436; PMID:85234394; PMID:3924897  
A:Accession: A24436  
A:Molecule type: DNA  
A:Residues: 1-549 <NAK>  
A:Cross-references: GB:M11450  
A:Experimental source: plasmid pAT5  
R:Jorgensen, P.L.; Hansen, C.K.; Poulsen, G.B.; Diderichsen, B.  
Gene 96, 37-41, 1990  
A:Title: In vivo genetic engineering: homologous recombination as a tool for plasmid con  
A:Reference number: I39772; PMID:91092499; PMID:2265757  
A:Accession: I39777  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-45 <RES>  
A:Cross-references: GB:M62638; NID:GI42514; PIDN:AAA2242.1; PID:GI42515  
C:Comment: Alpha-amylose genes have been found on plasmids and in multiple copies on the  
C:Genetics:  
A:Gene: amyS  
A:Genome: plasmid  
A:Start codon: GTG  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Superfamily: alpha-amylose, amyloliquefaciens type; alpha-amylose core homology  
C:Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac  
F:1-34/Domain: signal sequence #status predicted <Sig>  
F:35-549/Product: alpha-amylose #status experimental <MAT>  
F:235-368/Domain: alpha-amylose core homology <AMY>  
F:139, 237, 272/Binding site: calcium (Asp, Asp, His) #status predicted  
F:268, 298, 365/Active site: Asp, Glu, Asp #status predicted

Query Match 18.8%; Score 461.5; DB 1; Length 549;  
Best Local Similarity 28.4%; Pred. NO. 8.6e-26;  
Matches 146; Conservative 65; Mismatches 170; Indels 133; Gaps 25;

QY 11 GVIMQAFYDVPSSGIWDTTQKIPWYDAGISAIWIPPPASKMGAGYSGYDYPDFD 70

C;Date: 30-Jun-1998 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
 C;Accession: A24549; I39501; I39770  
 R;Gray, G.L.; Mainzer, S.E.; Rey, M.W.; Lamsa, M.H.; Kindie, K.L.; Carmona, C.; Requadt,  
 J. Bacteriol. 166, 635-643, 1986  
 A;Title: Structural genes encoding the thermophilic alpha-amylases of Bacillus stearother-  
 A;Reference number: A31817; MUID:86195857; PMID:3009417  
 A;Accession: A24549  
 A;Molecule type: DNA  
 A;Residues: 1-549 <GRA>  
 A;Cross-references: GB:M13255; NID:g142512; PIDN:AAA22241.1; PID:g142513  
 A;Experimental source: Genomic DNA of strain NZ-3  
 R;Satch, H.; Nishida, H.; Isono, K.  
 J. Bacteriol. 170, 1034-1040, 1988  
 A;Title: Evidence for movement of the alpha-amylase gene into two phylogenetically dista  
 A;Reference number: I39501; MUID:86139156; PMID:3257753  
 A;Accession: I39501  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 536-549 <RES>  
 A;Cross-references: GB:M29577; NID:g142476; PIDN:AAA22225.1; PID:g142478  
 A;Experimental source: strain DY-5  
 A;Accession: I39770  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 536-549 <RE2>  
 A;Cross-references: GB:M29578; NID:g142484; PIDN:AAA22228.1; PID:g142486  
 A;Experimental source: strain 799  
 C;Comment: Alpha-amylase genes have been found on plasmids and in multiple copies on the  
 C;Genetics:  
 A;Start codon: GTG  
 C;Function:  
 A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
 A;Pathway: glycogen/starch degradation  
 C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
 C;Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac  
 F;1-34/DNA: signal sequence #status predicted <SIG>  
 F;35-549/Product: alpha-amylase #status predicted <MAT>  
 F;235-368/DNA: alpha-amylase core homology <AMY>  
 F;139,237,272/Binding site: calcium (Asp, Asp, His) #status predicted  
 F;268,298,365/Active site: Asp, Glu, Asp #status predicted  
 Query Match 19.3%; Score 475.5; DB 1; Length 549;  
 Best Local Similarity 29.2%; Pred. No. 8.3e-27;  
 Matches 151; Conservative 61; Mismatches 166; Indels 139; Gaps 25;  
 QY 11 GVIMQAFYHVPSSGIGIWDITIRQKIPEDYDAGISAIWIPPSKMGVSGYDYPDFD 70  
 DB 40 GTMMQYFEWYLPDDGLTWTCKVANEANUSLGLTALWLPAYKTSRS-DVGYGYDYLDY 98  
 QY 71 LGEYDQKGVETFRFGSKQELVNMINTAHAYGMKVADIVINHRRAGD-LEW-----NPF 123  
 DB 99 LGEFNQKGVIRKYGTQKYLQAIQAHAAGQVYADVDFDKGAGDGTWDAVENP- 157  
 QY 124 VNDYTDTSKVASGYTANYLDHFENELHAGSGTFCGYPDICHKSNQY----- 175  
 DB 158 -----SDRNOEISGTY-----QIQAWTKDFNGRNGTYSFKRWYHFGVDWD 201  
 QY 176 -----WLWASQES-----YAA-----YLRSIG 192  
 DB 202 ESKLSRIYKFRIGKAWDEVDTEGNYDYLMDLDMDEPVTTELKNGWKVYVNTN 261  
 QY 193 IDAWRFDYKGYAPVYVVKWDLNWWG-----WAVEGYWDTWDAVLNWAYSSGA--KVF 244  
 DB 262 IDGFRDLAVKHIFKSPFPDLVSYSQTKPLFTVGEYSYDINKLHNYITKNGTMSLF 321  
 QY 245 DFALYK-----DEAFDNKIPALVASALQNGQTVSRDRPKAVTFVANHTD----- 292  
 DB 322 DAPLHNKFTYASKSGGAFF-----MSTLMN-NTLUNKOQPLAVTVNHDHTEFGQALQ 373  
 QY 293 ---IWNKYPAYAFILT-VEGQPTIFYRDE--EWLN-----KDKLKNLIWIHENTAGST 342  
 DB 374 SWDPWFKPLAYAFILTRQSGYPCVFYGYGIPQYNIPSLKSKIDPLLIARRDYAYG-T 432

QY 343 DIVYDNDDELI-FVRNGYGDKP--GLITVINLGSSKAGRWVYVPR-FAGACIHEYTNGLG 398  
 DB 433 QHDYLDHSDIILGWTEGVTKEKPGSLAALITDGP-GSKWYVQKHAGKVFYDLTGNS 491  
 QY 399 GWYDKYVSSGWVYLEAPAYDPANGQY-----GYSVW 430  
 DB 492 DTVT--INSDGW-----GFEKVGSGSVW 514  
 RESULT 5  
 ALSN  
 alpha-amylase (EC 3.2.1.1) precursor - Bacillus amyloliquefaciens  
 N;Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
 C;Species: Bacillus amyloliquefaciens  
 C;Date: 30-Nov-1980 #sequence\_revision 30-Jun-1987 #text\_change 18-Jun-1999  
 C;Accession: A92389; A90307; I39756; I39763; A00843  
 R;Takinen, K.; Pettersson, R.F.; Kalkinen, N.; Palva, I.; Soderlund, H.; Kaariainen, L.  
 J. Biol. Chem. 258, 1007-1013, 1983  
 A;Title: Amino acid sequence of alpha-amylase from Bacillus amyloliquefaciens deduced fr  
 A;Reference number: A92389; MUID:83108808; PMID:6185474  
 A;Contents: PUB110  
 A;Accession: A92389  
 A;Molecule type: DNA  
 A;Residues: 1-514 <TAK>  
 A;Cross-references: GB:J01542; GB:J01543; GB:M12033; GB:M12034; NID:g142428; PIDN:AAA2215  
 R;Chung, H.S.; Friedberg, F.  
 Biochem. J. 185, 387-395, 1980  
 A;Title: Sequence of the N-terminal half of Bacillus amyloliquefaciens alpha-amylase.  
 A;Reference number: A90307; MUID:80241725; PMID:6156671  
 A;Accession: A90307  
 A;Molecule type: protein  
 A;Residues: 32-53; 1', 55-63; 1', 65-78; 1', 80-83; 1', 85-222 <CHU>  
 R;Palva, I.; Pettersson, R.F.; Kalkinen, N.; Lehtovaara, P.; Sarvas, M.; Soderlund, H.;  
 Gene 15, 43-51, 1981  
 A;Title: Nucleotide sequence of the promoter and NH2-terminal signal peptide region of t  
 A;Reference number: I39756; MUID:82051236; PMID:6170539  
 A;Accession: I39756  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-96 <RES>  
 A;Cross-references: EMBL:V00092; NID:g39297; PIDN:CAA23430.1; PID:g39298  
 R;Ruchonen, L.; Hackman, P.; Lehtovaara, P.; Knowles, J.K.C.; Karaenen, S.  
 Gene 59, 161-170, 1987  
 A;Title: Efficient secretion of Bacillus amyloliquefaciens alpha-amylase cells by its ow  
 A;Reference number: I39763; MUID:88137952; PMID:2830166  
 A;Accession: I39763  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-39 <RE2>  
 A;Cross-references: GB:M18424; NID:g142430; PIDN:AAA22192.1; PID:g142431  
 C;Function:  
 A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
 A;Pathway: glycogen/starch degradation  
 C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
 C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation  
 F;1-31/DNA: signal sequence #status predicted <SIG>  
 F;32-514/Product: alpha-amylase #status predicted <MPT>  
 F;229-362/DNA: alpha-amylase core homology <AMY>  
 F;133,231,266/Binding site: calcium (Asn, Asp, His) #status predicted  
 F;262,292,359/Active site: Asp, Glu, Asp #status predicted  
 Query Match 19.2%; Score 472.5; DB 1; Length 514;  
 Best Local Similarity 29.9%; Pred. No. 1.3e-26;  
 Matches 147; Conservative 60; Mismatches 167; Indels 117; Gaps 22;  
 QY 11 GVIMQAFYHVPSSGIGIWDITIRQKIPEDYDAGISAIWIPPSKMGVSGYDYPDFD 70  
 DB 34 GTLMQYFEWYLPDDGLTWTCKVANEANUSLGLTALWLPAYKTSRS-DNGYGYDYLDY 92  
 QY 71 LGEYDQKGVETFRFGSKQELVNMINTAHAYGMKVADIVINHRRAGD-----LEWNP- 122  
 DB 93 LGEFOQKGVIRKYGTQKYLQAIQAHAAGQVYADVDFDKGAGDGTWDAVENP- 152

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:30:47 ; Search time 18 Seconds  
(without alignments)  
1261.255 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVMSYCGVG 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484.5	19.7	494	1	AMY2_SALTY
2	472.5	19.2	514	1	AMY_BACAM
3	461.5	18.6	549	1	AMY_BACST
4	458.5	18.6	512	1	AMY_BACLI
5	454.5	18.5	518	1	AMY6_BACS7
6	447.5	18.2	495	1	AMY2_ECOLI
7	400	16.3	421	1	AMYA_VIGMU
8	396.5	16.1	443	1	AM2A_ORYSA
9	390.5	15.9	445	1	AMC2_ORYSA
10	366.5	14.9	427	1	AMY2_HORVU
11	365.5	14.9	427	1	AMY1_HORVU
12	363.5	14.8	368	1	AMY3_HORVU
13	357.5	14.5	438	1	AM3B_ORYSA
14	353	14.4	437	1	AM3C_ORYSA
15	351.5	14.3	429	1	AMY6_HORVU
16	347.5	14.1	440	1	AM3A_ORYSA
17	347	14.1	428	1	AMY1_ORYSA
18	341.5	13.9	435	1	AM3D_ORYSA
19	336	13.7	437	1	AM3E_ORYSA
20	335.5	13.6	413	1	AMY3_WHEAT
21	284	11.5	548	1	AMT4_PEST
22	278	11.3	551	1	AMT4_PSSA
23	263.5	10.3	710	1	CDGT_THETU
24	242	9.8	719	1	AMY1_BACST
25	236	9.6	464	1	AMY1_AERYH
26	235.5	9.6	711	1	CDGT_BACST
27	232.5	9.5	712	1	CDGT_BACS3
28	230.5	9.4	718	1	CDGT_BACS5
29	230	9.4	713	1	CDGT_BACSP
30	229	9.3	713	1	CDG2_PAEWA
31	229	9.3	714	1	CDG1_PAEWA
32	227	9.2	704	1	CDGT_BACCH
33	226	9.2	713	1	CDGT_BACSO

34	224.5	9.1	718	1	CDGT_BACLI
35	223.5	9.1	718	1	CDGT_BACCI
36	214.5	8.7	713	1	CDGU_BACCI
37	211	8.6	919	1	AMY_STRLI
38	210	8.5	703	1	CDGT_BACS2
39	205	8.3	1196	1	AMYB_PASPO
40	201.5	8.2	713	1	CDGT_BACS8
41	201	8.2	1481	1	APU_THEET
42	200	8.1	581	1	AMYI_SCHPO
43	198.5	8.1	498	1	AMYB_ASAPW
44	198.5	8.1	499	1	AMYB_ASAPW
45	196.5	8.0	499	1	AMY_ASPSH

ALIGNMENTS

RESULT 1					
AMY2_SALTY					
ID	AMY2_SALTY	STANDARD;	PRT;	494 AA.	
AC	P26613;				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Cytoplasmic alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan				
DE	glucanohydrolase).				
GN	AMYA OR STM1963;				
OS	Salmonella typhimurium.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Salmonella.				
OX	NCBI_TaxID=602;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SW1103;				
RX	MEDLINE=93015717; PubMed=1400215;				
RA	Raha M., Kawagishi I., Mueller V., Kihara M., Macnab R.M.;				
RT	"Escherichia coli produces a cytoplasmic alpha-amylase, AmyA.";				
RL	J. Bacteriol. 174:6644-6652(1992).				
[2]					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=J72 / SGSC1412 / ATCC 700720;				
RX	MEDLINE=21534948; PubMed=11677609;				
RA	McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,				
RA	Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,				
RA	Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,				
RA	Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,				
RA	Waterston R., Wilson R.K.;				
RT	"Complete genome sequence of Salmonella enterica serovar Typhimurium				
RL	L72.";				
RL	Nature 413:852-856 (2001).				
[3]					
RP	SEQUENCE OF 1-6 FROM N.A.				
RC	STRAIN=J7M1103.				
RX	MEDLINE=92407478; PubMed=1527488;				
RA	Kawagishi I., Mueller V., Williams A.W., Irikura V.M., Macnab R.M.;				
RT	"Subdivision of flagellar region III of the Escherichia coli and				
RT	Salmonella typhimurium chromosomes and identification of two				
RT	additional flagellar genes.";				
RL	J. Gen. Microbiol. 138:1051-1065(1992).				
[4]					
RP	SEQUENCE OF 476-494 FROM N.A.				
RX	MEDLINE=93381452; PubMed=8371104;				
RA	Raha M., Kihara M., Kawagishi I., Macnab R.M.;				
RT	"Organization of the Escherichia coli and Salmonella typhimurium				
RT	chromosomes between flagellar regions IIIa and IIb, including a				
RT	large non-coding region.";				
RL	J. Gen. Microbiol. 139:1401-1407(1993).				
CC	-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic				
CC	linkages in oligosaccharides and polysaccharides.				
CC	-!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).				
CC	-!- SUBUNIT: Monomer (By similarity).				
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	-!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.				





DR PDB: 1E43; 21-JUN-01.  
DR InterPro: IPR006589; Alp\_aml\_cat\_sub.  
DR InterPro: IPR006047; Alp\_aml\_cat.  
DR InterPro: IPR006046; Glyco\_hydro\_13.  
DR Pfam: PF00128; alpha-amylase; 1.  
DR PRINTS: PR00110; ALPHAAMYLASE.  
DR SMART: SM00642; Aamy; 1.  
KW Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
KW Signal; 3D-structure  
FT SIGNAL 1 31  
FT CHAIN 32 514  
FT ACT\_SITE 262 262  
FT ACT\_SITE 292 292  
FT ACT\_SITE 359 359  
FT METAL 133 133  
FT METAL 190 190  
FT METAL 212 212  
FT METAL 214 214  
FT METAL 225 225  
FT METAL 231 231  
FT METAL 233 233  
FT METAL 235 235  
FT METAL 266 266  
FT METAL 331 331  
FT METAL 438 438  
FT METAL 461 461  
FT CONFLICT 54 54  
FT CONFLICT 64 64  
FT CONFLICT 79 79  
FT CONFLICT 84 84  
SQ SEQUENCE 514 AA; 58403 MW; 3D566B3FB5CDE7E CRC64;  
Query Match 19.2%; Score 472.5; DB 1; Length 514;  
Best Local Similarity 29.9%; Pred. No. 1.3e-26;  
Matches 147; Conservative 60; Mismatches 167; Indels 117; Gaps 22;  
QY 11 GVIMQAFVDPSPGIGIWDITRQKIPEDYDAGISAIWIPPPASKMGAYSGYDPYDFD 70  
DB 34 GTLMQYFEWYFNDQHQWKRLLQDAEHLSDIGITAVWIPPPAYKGLSQS-DNGYGPYDLYD 92  
QY 71 LGEYDQKTVETRFSGKQELVNMINTAHAYKVIADIVINHRAGD-----LEWNP- 122  
DB 93 LGEFOQKTVITKYTKSELQDAIGSLHSRNVQVGVVLMHKGADATEDVTAVEVNP 152  
QY 123 -----FVNDY---TWTFDSKVASGKYTYNYLDFFHFNELH----- 153  
DB 153 NNNQETSEYQIKAWTDFRFGNGN---TYSDFKWHVHPDGDWDSRKRISRFKPERGE 209  
QY 154 -----AGSGTGF--GYDICH-----KSHDQYWLWASQESYAAVLSIGIDA 195  
DB 210 GKANDWEVSSNGYDLYMDYADVDYDHPDVVAETKKG-----YANELSLDG 258  
QY 196 WRFDYVKGYPWVKDWL-----NWMGMAVGVEYWDVNTVDVNLW--AYSSGAKVDFEA 247  
DB 259 FRIDAAKHKFSLFDWVQAVRQATKEMFTVAIYQWQNNAGKLENYLNKTSFNQSVFDP 318  
QY 248 LYYKDEA-----FDKNIPALVSLONGTQVVSRRDPFKAVTFVANEHDT----- 292  
DB 319 LHFNLQASSGGGYDMR-----LLDG-TVVSRHPEKAVTFVENEHDTQPGQSLESTV 370  
QY 293 IIMNKYPAYAFILTYE-GOPTIFYRDY-----EELNKKLKNLIWIHENLAGGST 342  
DB 371 QTWFKPLAYAFILTRSGYPOVFDGMVYKGTGPKPIPSLKNIEPILKARKEYAYGPQ 430  
QY 343 DIVYDNDLEI-FVRNGYD--KPLGITYINLGSSKAGRWVYVYVFKFAGACIHEVTNGLG 399  
DB 431 H-DVIDHPDVIGTWREGDSSAAKGLAALITDGGGSKRVYAGLKNAGETWYDITGNRSD 489  
QY 400 WVDKVVSSGW 410  
DB 490 TVK--IGSDGW 498

RESULT 3  
AMY\_BACST  
ID - AMY\_BACST STANDARD; PRT; 549 AA.  
AC P06279; O45519;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
glucanohydrolase).  
GN AMYS.  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
OX NCBI\_TaxID=1422;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-39.  
RX MEDLINE=85234394; PubMed=3924897;  
RA Nakajima R., Imanaka T., Aiba S.;  
RT "Nucleotide sequence of the Bacillus stearothermophilus alpha-amylase  
gene.";  
RL J. Bacteriol. 163:401-406 (1985).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX STRAIN=DV5/PHI300;  
RA Ihara H., Sasaki T., Tsuboi A., Yamagata H., Tsukagoshi N., Ueda S.;  
RT "Complete nucleotide sequence of a thermophilic alpha-amylase gene:  
homology between prokaryotic and eukaryotic alpha-amylases at the  
active sites.";  
RL J. Biochem. 98:95-103 (1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=NZ-3;  
RA Gray G.L., Mainzer S.B., Rey M.W., Lamsa M.H., Kindle K.L.,  
RA Carmona C., Requa C.;  
RT "Structural genes encoding the thermophilic alpha-amylases of  
Bacillus stearothermophilus and Bacillus licheniformis.";  
RL J. Bacteriol. 166:635-643 (1986).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Suominen I., Karp M., Lautamo J., Knowles J., Mantea P.;  
RT "Thermostable alpha amylase of Bacillus stearothermophilus: cloning,  
expression, and secretion by Escherichia coli.";  
RL (In) Chaloupka J., Krumphanz V. (eds.);  
RL Extracellular enzymes of microorganisms, pp.129-137, Plenum Press,  
New York (1987).  
RN [5]  
RP SEQUENCE OF 1-122 FROM N.A., AND SEQUENCE OF 35-48.  
RX STRAIN=DY-5;  
RA MEDLINE=86059211; PubMed=2999073;  
RA Tsukagoshi N., Iritani S., Sasaki T., Takemura T., Ihara H.,  
RA Igota Y., Yamagata H., Ueda S.;  
RT "Efficient synthesis and secretion of a thermophilic alpha-amylase by  
protein-producing Bacillus brevis 47 carrying the Bacillus  
stearothermophilus amylase gene.";  
RL J. Bacteriol. 164:1182-1187 (1985).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=21125602; PubMed=11226887;  
RA Savd D., Fujimoto Z., Takase K., Matsumura M., Mizuno H.;  
RT "Crystal structure of Bacillus stearothermophilus alpha-amylase:  
possible factors determining the thermostability.";  
RL J. Biochem. 129:461-468 (2001).  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
linkages in oligosaccharides and polysaccharides.  
CC -!- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.  
CC -!- SIMILARITY: Monomer.  
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; M11450; ARA22235-2; -  
 DR EMBL; X02769; CRA26547-1; -  
 DR EMBL; M57457; ARA22227-1; -  
 DR EMBL; M13255; ARA22241-1; -  
 DR PIR; A24436; A24436.  
 DR PIR; A91999; ALBSF.  
 DR PDB; 1HWX; 05-AUG-03.  
 DR InterPro; IPR006589; Alp\_aml cat sub.  
 DR InterPro; IPR006047; Alpha\_aml cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amyrase; 1.  
 DR PRINTS; PR00110; ALPHAAAMYLAASE.  
 DR SMART; SM00642; Aamy; 1.  
 DR Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 KW Signal; 3D-structure.  
 FT SIGNAL 1 34  
 FT CHAIN 35 549  
 FT ACT\_SITE 268 268  
 FT ACT\_SITE 272 272  
 FT ACT\_SITE 365 365  
 FT METAL 139 139  
 FT METAL 196 196  
 FT METAL 218 218  
 FT METAL 220 220  
 FT METAL 231 231  
 FT METAL 237 237  
 FT METAL 238 238  
 FT METAL 239 239  
 FT METAL 272 272  
 FT METAL 337 337  
 FT METAL 339 339  
 FT METAL 440 440  
 FT METAL 441 441  
 FT METAL 464 464  
 FT CONFLICT 13 13  
 FT CONFLICT 19 19  
 FT CONFLICT 23 23  
 FT CONFLICT 31 31  
 FT CONFLICT 107 107  
 FT CONFLICT 167 167  
 FT CONFLICT 179 179  
 FT CONFLICT 251 251  
 FT CONFLICT 260 260  
 FT CONFLICT 284 284  
 FT CONFLICT 312 312  
 FT CONFLICT 338 338  
 FT CONFLICT 342 342  
 FT CONFLICT 346 346  
 FT CONFLICT 376 376  
 FT CONFLICT 526 526  
 FT CONFLICT 527 527  
 FT CONFLICT 535 535  
 SQ SEQUENCE 549 AA; 62670 MW; 3A2D93A955E79D3 CRC64;  
 Query Match 18.8%; Score 461.5; DB 1; Length 549;  
 Best Local Similarity 28.4%; Pred. No. 8.6e-26;  
 Matches 146; Conservative 65; Mismatches 170; Indels 133; Gaps 25;  
 QY 11 GVINQAFVWVPSGGIWDITIRKIPEDWDAGISAIWIPPSKMGAGSMGVDPYDFPD 70  
 DB 40 GTMMQCFEYLPDDGTLTKVANEANNLSLIGITLWLPAYKGTGRS-DVGVGVDLYD 98  
 QY 71 LGEYDQKTVETRGSKQELVNMINTAHAYGMKVIADIVINHEAGD-LGW-----NPF 123  
 DB 99 LGFENQKGAETKTKYTAQVLAQIAAHAGMGMVYADVDFDHKGAGDGTWDAVEVNP 158  
 QY 124 VND-----YTWTFDSKVASGKYTANYLDLPHNELHAG----- 155  
 DB 159 DRNQEISGTQIAQWTKFDFFGRGN---TYSSPKRWYHFDGVDWDESKRLSRIVKFRGI 215

QY 156 -----DSGTFG-----GYDICH-D-KSWDQYWLWASQESYAAYLRSIGIDA 195  
 DB 216 GKAWDEVDTEGNYDYLMDYADLDMDHPVTVTELKSGKW-----YVNTNIDG 264  
 QY 196 WRDYKGYAPVWVKQKLNWGG-----WAYGEYDVTNVDVAVLWAYSSGA--KVFDA 247  
 DB 265 FRLDAVAKHIFSPFDWLSVRSQTKPLFTVGEYWSYDINKLHNYIMKNTGMTSLFDAP 324  
 QY 248 LYYKMDA-----FDKNIPALVSALQNGQTVSSRDPFKAVTFVANHDT----- 292  
 DB 325 LHNKFYTASKSGGTDFWRTLMT-----NTLMKQDPTLAVTFVDNHDTEPGQALQSWV 376  
 QY 293 IINWKPAYAFILT-YEGOPTIEYRDE--EWN-----KDKLKLNIWIHNLGAGSTDIV 345  
 DB 377 DPFKPLAYAFILTREQEYPCVFIYGIYQNPVIFSLASKIDPLLIARDVAYG-TQHD 435  
 QY 346 YYDNELI-FVRNGYDKP--GLITYINILGSSKAGRWVYVK-FAGACHEYTNLGGWV 401  
 DB 436 YLDHSDIIGWTREGVTEKPGSLAALITDQPG-GSKWMTYVQKHAGKVFYDLTGNSDTV 494  
 QY 402 DKYVYSSGVYLEAPAYDPAFGY-----GYSVW 430  
 DB 495 T--INSDG-----GEFKVNGGSVSW 514  
 RESULT 4  
 ID AMY\_BACLI STANDARD; PRT; 512 AA.  
 AC P02778; Q84171;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Alpha-amyrase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE Glucanohydrolase) (BLA).  
 GN AMYS OR AMYL.  
 OS Bacillus licheniformis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1402;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27811;  
 RX MEDLINE=86111694; PubMed=2418011;  
 RA Yuuki T., Nomura T., Tezuka H., Tsuboi A., Yamagata H.,  
 RA Tsukagoshi N., Ueda S.;  
 RT "Complete nucleotide sequence of a gene coding for heat- and  
 RT pH-stable alpha-amyrase of Bacillus licheniformis: comparison of the  
 RT amino acid sequences of three bacterial liquefying alpha-amyrases  
 RT deduced from the DNA sequences."  
 RL J. Biochem. 98:1147-1156(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86195857; PubMed=3009417;  
 RA Gray G.L., Mainzer S.B., Rey M.W., Lamsa M.H., Kindle K.L.,  
 RA Carmona C., Requaet C.;  
 RT "Structural genes encoding the thermophilic alpha-amyases of  
 RT Bacillus stearotherophilus and Bacillus licheniformis."  
 RL J. Bacteriol. 166:635-643(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Shahhoseini M., Ziaei A.A., Ghaemi N., Pourbabaie A.A.;  
 RT "An unusual DNA sequence encoded a hyperthermostable alpha-amyase."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-104 FROM N.A.  
 RX MEDLINE=84185455; PubMed=6609154;  
 RA Stephens M.A., Ortlepp S.A., Ollington J.F., McConnell D.J.;  
 RT "Nucleotide sequence of the 5' region of the Bacillus licheniformis  
 RT alpha-amyase gene: comparison with the B. amyloliquefaciens gene."  
 RL J. Bacteriol. 158:369-372(1984).  
 RN [5]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=89213924; PubMed=2540150;

RA Laoid B.M., Chambliss G.H., McConnell D.J.;  
 RT "Bacillus licheniformis alpha-amylase gene, amyL, is subject to  
 RT promoter-independent catabolite repression in *Bacillus subtilis*.";   
 RL J. Bacteriol. 171:2435-2442(1989).  
 RN [6]  
 RP SEQUENCE OF 30-47.  
 RX MEDLINE=82098050; PubMed=6172418;  
 RA Kuhn H., Fietzek P.P., Lampen J.O.;  
 RT "N-terminal amino acid sequence of *Bacillus licheniformis*  
 RT alpha-amylase: comparison with *Bacillus amyloliquefaciens* and  
 RL *Bacillus subtilis* enzymes.";   
 RL J. Bacteriol. 149:372-373(1982).  
 RN [7]  
 RP MAPPING OF SUBSTRATE-BINDING SITE.  
 RX MEDLINE=2192788; PubMed=11997021;  
 RA Kandra L., Gyemant G., Remenyik J., Hovanszki G., Liptak A.;  
 RT "Action pattern and substrate mapping of *Bacillus licheniformis*  
 RT alpha-amylase (BLA) with modified maltooligosaccharide substrates.";   
 RL FEBS Lett. 518:79-82(2002).  
 RN [8]  
 RP MUTAGENESIS OF HIS-64; HIS-162; HIS-276; HIS-322; HIS-435 AND HIS-479.  
 RX STRAIN=ATCC 6598;  
 RA Declerck N., Joyet P., Gaillardin C., Maasson J.M.;  
 RT "Use of amber suppressors to investigate the thermostability of  
 RT *Bacillus licheniformis* alpha-amylase. Amino acid replacements at 6  
 RT histidine residues reveal a critical position at His-133.";   
 RL J. Biol. Chem. 265:15481-15488(1990).  
 RN [9]  
 RP MUTAGENESIS OF ALA-238.  
 RX STRAIN=ATCC 6598;  
 RA Declerck N., Joyet P., Trosset J.Y., Garnier J., Gaillardin C.;  
 RT "Hyperthermostable mutants of *Bacillus licheniformis* alpha-amylase:  
 RT multiple amino acid replacements and molecular modelling.";   
 RL Protein Eng. 8:1029-1037(1995).  
 RN [10]  
 RP MUTAGENESIS OF ASP-150; ASN-155; ARG-175; ASP-193; ASN-201; GLN-207;  
 RP ASN-217; ASN-219; ASN-221; ASP-229; ASP-233; ALA-298; GLU-300; GLN-359  
 RP AND GLU-365.  
 RX STRAIN=ATCC 6598;  
 RA Declerck N., Machius M., Wiegand G., Huber R., Gaillardin C.;  
 RT "Probing structural determinants specifying high thermostability in  
 RT *Bacillus licheniformis* alpha-amylase.";   
 RL J. Mol. Biol. 301:1041-1057(2000).  
 RN [11]  
 RP MUTAGENESIS OF GLN-293 AND ASN-294.  
 RX STRAIN=ATCC 6598;  
 RA Declerck N., Machius M., Joyet P., Wiegand G., Huber R.,  
 RT Gaillardin C.;  
 RT "Hyperthermostabilization of *Bacillus licheniformis* alpha-amylase and  
 RT modulation of its stability over a 50 degrees C temperature range.";   
 RL Protein Eng. 16:287-293(2003).  
 RN [12]  
 RP MUTAGENESIS OF TRP-292 AND VAL-315.  
 RX STRAIN=ATCC 27811;  
 RA Declerck N., Machius M., Wiegand G., Huber R.,  
 RT Rivera M.H., Lopez-Munguia A., Soberon X., Saab-Rincon G.;  
 RT "Alpha-amylase from *Bacillus licheniformis* mutants near to the  
 RT catalytic site: effects on hydrolytic and transglycosylation  
 RT activity.";   
 RL Protein Eng. 16:505-514(2003).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX STRAIN=ATCC 27811;  
 RA Machius M., Wiegand G., Huber R.;  
 RT "Crystal structure of calcium-depleted *Bacillus licheniformis* alpha-  
 RT amylase at 2.2-A resolution.";   
 RL J. Mol. Biol. 246:545-559(1995).  
 RN [14]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=98212915; PubMed=9551551;  
 RA Machius M., Declerck N., Huber R., Wiegand G.;  
 RT "Activation of *Bacillus licheniformis* alpha-amylase through a  
 RT disorder-to-order transition of the substrate-binding site mediated  
 RT by a calcium-sodium-calcium metal triad.";   
 RL Structure 6:281-292(1998).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 330-512.  
 RX MEDLINE=20384196; PubMed=10924103;  
 RA Brzozowski A.M., Lawson D.M., Turkenburg J.P., Bisgaard-Prantzen H.,  
 RT Svendsen A., Borchert T.V., Dauter Z., Wilson K.S., Davies G.J.;  
 RT "Structural analysis of a chimeric bacterial alpha-amylase.  
 RT High-resolution analysis of native and ligand complexes.";   
 RL Biochemistry 39:9099-9107(2000).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF MUTANT  
 RX H162V/N219F/A238V/Q293S/N294Y.  
 RA STRAIN=ATCC 6598;  
 RX MEDLINE=22538505; PubMed=12540849;  
 RA Machius M., Declerck N., Huber R., Wiegand G.;  
 RT "Kinetic stabilization of *Bacillus licheniformis* alpha-amylase through  
 RT introduction of hydrophobic residues at the surface.";   
 RL J. Biol. Chem. 278:11546-11553(2003).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- BIOTECHNOLOGY: Used in the food industry for high temperature  
 CC liquefaction of starch-containing mash and in the detergent  
 CC industry to remove starch. Sold under the name Termamyl by  
 CC Novozymes.  
 CC -1- MISCELLANEOUS: Able to work at relatively high (alkaline) pH  
 CC values (up to pH 11) and at high temperatures (up to 100 degrees  
 CC Celsius).  
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X03236; CAA26981.1; -  
 CC EMBL; M38570; AAA22226.1; -  
 CC EMBL; M32556; AAA22240.1; -  
 CC EMBL; K01984; AAA22193.1; -  
 CC EMBL; AF438149; AAO26743.1; -  
 CC EMBL; M26412; AAA22237.1; -  
 CC EMBL; A17930; CAA01355.1; -  
 CC PIR; A91997; AUBSL.  
 CC DR PDB; 1BLI; 23-MAR-99.  
 CC DR PDB; 1BPL; 17-AUG-96.  
 CC DR PDB; 1E3X; 21-JUN-01.  
 CC DR PDB; 1E3Z; 24-JUN-03.  
 CC DR PDB; 1E40; 24-JUN-03.  
 CC DR PDB; 1E43; 21-JUN-01.  
 CC DR PDB; 1O80; 03-APR-03.  
 CC DR PDB; 1VUS; 12-MAR-97.  
 CC DR InterPro; IPR006589; Alp\_aml\_cat\_sub.  
 CC DR InterPro; IPR006047; Alpha\_aml\_cat.  
 CC DR InterPro; IPR006046; Glyco\_hydro\_13.  
 CC DR Pfam; PF00128; alpha-amylase; 1.  
 CC DR PRINTS; PR00110; ALPHAMYLASE.  
 CC DR SMART; SM00642; Amy; 1.  
 CC DR Signal; 3D-structure.  
 CC KW Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 FT SIGNAL 1 29  
 FT CHAIN 30 512 ALPHA-AMYLASE.  
 FT ACT\_SITE 260 260  
 FT ACT\_SITE 264 264

FT ACT SITE 357 357  
 FT METAL 133 133  
 FT METAL 190 190  
 FT METAL 210 210  
 FT METAL 212 212  
 FT METAL 223 223  
 FT METAL 229 229  
 FT METAL 231 231  
 FT METAL 233 233

Query Match 18.6%; Score 458.5; DB 1; Length 512;  
 Best Local Similarity 28.6%; Pred. No. 1.3e-25;  
 Matches 139; Conservative 68; Mismatches 170; Indels 109; Gaps 17;

QY 11 GVIMQAFYWDVPSGGIWMWDTIRQKIPWYDAGISAIWIPPPASKMGAYSMGYDYPDFD 70  
 Db 34 GTLMQYFEWYMPDQHWKRLQNDLSAELHGTAVMPPAYKGTQA-DVGYGAYDLYD 92

QY 71 LGEYDQKGTETRFSGKQELVNMINTAHAYCMKVIADIVLNHRAGGLEWNPVNDYTW 130  
 Db 93 LGEFHQKGTETRFSGKQELVNMINTAHAYCMKVIADIVLNHRAGGLEWNPVNDYTW 152

QY 131 DFSKVASGKYT-ANYLDFHPNELH- 153  
 Db 153 DNRNVISGEHRIKAWTHFPPGRGSTYSDFKWHYHFDGTDWDSRKLRIYKFGKAWD 212

QY 154 --AGDSTFG--CYPDICH- --KSDQYWLWASQESYAAAYLSIGIDAWRPD 200  
 Db 213 WEVSNENGYDLYMADIDYDHPDVAEIKRWGTW- --YANELQDGRFLDA 261

QY 201 VKGYAPWVYKDLNWWGG- --WAYGEYDNTNDAVLNWAYSS- --GAKVDFEALYIKM 252  
 Db 262 VKHIFSELRDWNHVRKTEKGMFTVAEYKQNDLGALENYLNTFNHVSFDPVPLHYQF 321

QY 253 DEA- --FNKNI PALVSALQNGQVVSERDPKAVTFVANHTD- --IINNK 297  
 Db 322 HAASQGGYDMRKL- --LNTSVSKHPLKAVTFVDNHDTPQGSLSSTVOTWPK 373

QY 298 YPAAVAFILTYE-GOFTIFYRDY- --EWLNRKDLKNIWIHNLAGSDTDIVY 347  
 Db 374 PLAYAFILTRSGYPOVYFGDMYTKGDSQREIPALKHKEIPILKARKQYVAGQH-DYF 432

QY 348 DNEDEL-FVRNGYGD- --KPLGITVINLSSKAGRWVYVPEACACHETVGNLGGWVYK 404  
 Db 433 DHDIVGTREGDSSVANSGLAALITDGPQGAIRMYVGRQNAGETWHDITGNRSEPV-V 490

QY 405 VYSGW 410  
 Db 491 INSEGW 496

RESULT 5  
 AMT6 BACS7  
 ID AMT6 BACS7 STANDARD; PRT; 518 AA.  
 AC P19571;  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Glucan 1,4-alpha-maltohexaosidase precursor (EC 3.2.1.98) (G6-amyase)  
 DE (Maltotetraose-producing amyase) (Exo-maltohexaosidase).  
 OS Bacillus sp. (strain 707).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1416;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 34-36.  
 RX MEDLINE=88162814; PubMed=3258152;  
 RA Tsukamoto A., Kimura K., Ishii Y., Takano T., Yamane K.;  
 RT "Nucleotide sequence of the maltotetraose-producing amyase gene from  
 an alkalophilic Bacillus sp. #707 and structural similarity to  
 RL liquefying type alpha-amyases."  
 RL Biochem. Biophys. Res. Commun. 151:25-31(1988).  
 CC -/- CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-D-glucosidic linkages  
 CC in amylaceous polysaccharides so as to remove successive

CC maltotetraose residues from the non-reducing chain ends.  
 CC -/- COPACTOR: Binds 2 calcium ions and 1 sodium ion per subunit (By  
 CC similarity).  
 CC -/- PATHWAY: Starch degradation.  
 CC -/- SUBCELLULAR LOCATION: Secreted.  
 CC -/- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M18862; AAA22231.1; -;  
 CC PIR; A27705; A27705.  
 CC HSP; P06278; IVJS.  
 CC InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
 CC InterPro; IPR006047; Alpha\_amy1\_cat.  
 CC InterPro; IPR006046; Glyco\_Hydro\_13.  
 CC Pfam; PF00128; alpha-amyase; 1.  
 CC PRINTS; PR00110; ALPHAAMYLASE.  
 CC SMART; SM00642; Amy; 1.  
 CC Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal.  
 CC SIGNAL 1 33  
 CC CHAIN 34 518 GLUCAN 1,4-ALPHA-MALTOHEXAOSIDASE.  
 CC ACT\_SITE 269 269 BY SIMILARITY.  
 CC ACT\_SITE 273 273 BY SIMILARITY.  
 CC ACT\_SITE 366 366 BY SIMILARITY.  
 CC METAL 139 139 CALCIUM 1 (BY SIMILARITY).  
 CC METAL 196 196 CALCIUM 2 AND SODIUM (BY SIMILARITY).  
 CC METAL 219 219 CALCIUM 2 (VIA CARBONYL OXYGEN) (BY  
 CC SIMILARITY).  
 CC METAL 221 221 CALCIUM 2 AND SODIUM (BY SIMILARITY).  
 CC METAL 232 232 CALCIUM 1 AND SODIUM (BY SIMILARITY).  
 CC METAL 238 238 CALCIUM 1 AND SODIUM (BY SIMILARITY).  
 CC METAL 240 240 CALCIUM 2 (BY SIMILARITY).  
 CC METAL 242 242 CALCIUM 2 (BY SIMILARITY).  
 CC METAL 273 273 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY  
 CC SIMILARITY).  
 CC SEQUENCE 518 AA; 59009 MW; 3A961E21612682C4 CRC64;

Query Match 18.5%; Score 454.5; DB 1; Length 518;  
 Best Local Similarity 29.1%; Pred. No. 2.5e-25;  
 Matches 141; Conservative 71; Mismatches 166; Indels 107; Gaps 23;

QY 11 GVIMQAFYWDVPSGGIWMWDTIRQKIPWYDAGISAIWIPPPASKMGAYSMGYDYPDFD 70  
 Db 40 GTLMQYFEWYMPDQHWKRLQNDLSAELHGTAVMPPAYKGTQA-DVGYGAYDLYD 98

QY 71 LGEYDQKGTETRFSGKQELVNMINTAHAYCMKVIADIVLNHRAGD- --LEWNP- 122  
 Db 99 LGEFNQKGTETRFSGKQELVNMINTAHAYCMKVIADIVLNHRAGD- --LEWNP- 158

QY 123 - --FVNDY- --WTFDFSKVASGK- --YTANYLDFHPNELHAGSDTGFEGYD 165  
 Db 159 NRNQEVTEGTIEAWTRFDPPGRGNTHSSFKWYHFDGVDWDQSRLLNNRIYKFRG- -- 215  
 QY 166 ICHDKSWD- --QYWLWAS- --QESYAAAYLSIGIDAWRPDYK 203  
 Db 216 - --HGRKAWDEVDTENGYDYLMDHDPPEVNLNRNWGVYNTLGLDGFRIADYKH 273

QY 204 YAPWVYKDLNWWGG- --WAVGEYDNTNDAV- --LNWAYSSGAKVDFEALYK 251  
 Db 274 IKYSFTEDLNHVSATGKMFVAEFAEKNLQKTNWNS- --VFDVPLHYN 329

QY 252 MDEA- --FDKNI PALVSALQNGQVVSERDPKAVTFVANHTD- --IINNK 296  
 Db 330 LYNAKSGGNYDMRNT- --FNG-TVQSHPHSHAVTFVDNHDSDQPEALESEVEWF 381

QY 297 KYPAAVAFILTYE-GOFTIFYRDYEW- --LNKDKLKNLIWIHNLAGSDTDIVYDN 349



QY 337 LAGSTDIYVNDDEL-FVRNGYGDKPLITYINLG 372  
DB 405 FAHQ-VQTLFFDHPNCIAFSRSGTDFEFCGVVMSNG 440

RESULT 7

AMYA\_VIGMU  
ID AMYA\_VIGMU STANDARD; PRT; 421 AA.  
AC P17859;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
glucanohydrolase).  
DE AMY1.1.  
GN Vigna mungo (rice bean) (Black gram).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurasids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.  
OX NCBI\_TaxID=3915;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Cotyledon;  
RC MEDLINE=90332425; PubMed=2377468;  
RA Yamauchi D., Minamikawa T.;  
RT "Nucleotide sequence of cDNA for alpha-amylase from cotyledons of  
germinating Vigna mungo seeds".  
RL Nucleic Acids Res. 18:4250-4250(1990).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=94120017; PubMed=8290640;  
RA Takeuchi H., Yamauchi D., Wada S., Minamikawa T.;  
RT "Nucleotide sequence of the alpha-amylase gene from Vigna mungo";  
PL Plant Physiol. 103:1459-1459(1993).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
linkages in oligosaccharides and polysaccharides.  
CC -1- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53049; CAA37217.1; -;  
CC EMBL; X73301; CAA51734.1; -;  
CC PIR; S10514; S10514.  
CC HSP; P04063; IAVA.  
DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
DR InterPro; IPR006047; Alp\_amy1\_cat.  
DR InterPro; IPR006046; Glyco\_hydro\_13.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00110; ALPHAMYLASE.  
DR SMART; SM00642; Amy; 1.  
DR Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
KW Signal.  
FT SIGNAL 1 23 PROBABLE.  
FT CHAIN 24 421 ALPHA-AMYLASE.  
FT ACT\_SITE 201 201 BY SIMILARITY.  
FT ACT\_SITE 309 309 BY SIMILARITY.  
FT METAL 113 113 CALCIUM 1 (BY SIMILARITY).  
FT METAL 130 130 CALCIUM 2 (BY SIMILARITY).  
FT METAL 133 133 CALCIUM 2 (BY SIMILARITY).  
FT METAL 135 135 CALCIUM 2 (BY SIMILARITY).  
FT METAL 139 139 CALCIUM 3 (BY SIMILARITY).  
FT METAL 149 149 CALCIUM 3 (BY SIMILARITY).  
FT METAL 160 160 CALCIUM 3 (BY SIMILARITY).  
FT METAL 168 168 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY  
SIMILARITY).

FT METAL 170 170 CALCIUM 1 AND 3 (BY SIMILARITY).  
SQ SEQUENCE 421 AA; 46888 MW; 15CA0DABA3DB4656 CRC64;  
Query Match 16.3%; Score 400; DB 1; Length 421;  
Best Local Similarity 28.6%; Pred. No. 1.6e-21;  
Matches 118; Conservative 57; Mismatches 109; Indels 128; Gaps 19;  
QY 12 VIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGGMGAYSMGYDYPDFDL 71  
DB 25 LIFQGFNWESSKGGWNSIKNSIPDLANAGITHVWLPSPQSVS---PEGYLPGRLYDL 81  
QY 72 GEYDQKGTVEFGSQKQELVNMINTAHVMKVADIIVINHEA-----GGD 117  
DB 82 D-----ASKYGSKNELSLIAFHEKGIKCLADIVINHRTAERKOGRGYICFEGT 133  
QY 118 ----LEWNPVNDYTWTFDSKVASGKYTANYLDFHFNELHAGDSGTFGY---PDICH-- 168  
DB 134 PDSRQMDGP---SFICRDTAYSDG--TGN-----NDSGE--GYDAAPDIDHLN 175  
QY 169 ---DKSMDQYWLWASQESVAAVLSIGIDARPDYKGYAPVVKWLNW--GGWAVGEY 224  
DB 176 PQVQRELSWMLKTE-----IGFDGWRDFVKGYAPSKITMSEQTKPDFAVGSK 227  
QY 225 WDT-----NVD-----AVLNWAYSSGAKVDFALVYKMDAEPDNKNIPALVSALQ 269  
DB 228 WDSISYQDGKPNYNQSHRGALVNVVESAGGAI-----TAFDFTTKGILQANVO 277  
QY 270 -----NGQT--VVSRRDPFKAVTFVANHTDIIWNKYP-----AYAFILTYEGQ 310  
DB 278 GELWRLIDPNGKPGMIGVKNPENAFTIDNHTGSTQRLWPPSPDKVMQGYAVILTHPGT 337  
QY 311 PTIPRYDEWLKDKKLNLIWHENLAGSTDIVYDNDLIFVNGYGDK 362  
DB 338 PSIFYDFHFDWGLKEQIAKL-----SSIRLRNGINEK 369

RESULT 8

AMZA\_ORYSA  
ID AMZA\_ORYSA STANDARD; PRT; 443 AA.  
AC P27935;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Alpha-amylase isozyme 2A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
glucanohydrolase).  
GN AMY1.5 OR AMY2A.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=cv. Japonica M202;  
RX MEDLINE=92175526; PubMed=1541400;  
RA Huang N., Reini S.J., Rodriguez R.L.;  
RT "AMY2A, a novel alpha-amylase-encoding gene in rice.";  
RL Gene 111:223-228(1992).  
CC -1- FUNCTION: Important for breakdown of endosperm starch during  
germination.  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
linkages in oligosaccharides and polysaccharides.  
CC -1- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
CC -1- SUBUNIT: Monomer.  
CC -1- DEVELOPMENTAL STAGE: Expressed at a high level during germination  
in the aleurones cells under the control of the plant hormone  
gibberellic acid and in the developing grains at a low level.  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M741177; AAA33894.1; --  
 CC PIR; JQ1527; JQ1527.  
 CC HSSP; P04063; IAVA.  
 CC Gramene; P27935; --  
 DR InterPro; IPR006047; Alpha\_aml1\_cat.  
 DR Pfam; P00128; alpha-amy1ase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 KW Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 21  
 FT CHAIN 22 443  
 FT ACT\_SITE 202 202  
 FT ACT\_SITE 314 314  
 FT METAL 113 113  
 FT METAL 130 130  
 FT METAL 133 133  
 FT METAL 135 135  
 FT METAL 139 139  
 FT METAL 149 149  
 FT METAL 160 160  
 FT METAL 163 163  
 FT METAL 164 164  
 FT METAL 165 165  
 FT METAL 168 168  
 FT METAL 170 170  
 FT SEQUENCE 443 AA; 48527 MW; 7B3F9264404F67F6 CRC64;  
 Query Match 16.1%; Score 396.5; DB 1; Length 443;  
 Best Local Similarity 26.4%; Pred. No. 3e-21;  
 Matches 126; Conservative 67; Mismatches 150; Indels 135; Gaps 21;  
 QY 12 VIMQAFYDVV--PSGGIWDITROKIPEDWDAGISAIWIPSPKMGAGVSMGYDPYDF 69  
 Db 24 ILFGQFNWESRQSGG--WYNLMGKVDVIAAGVTHVWLPSPHSVS---TQGWMPGR 79  
 QY 70 DLGEYDQGTGTVETFRGSKQELVNMINTAHAYGMKVIADIVINHRAG----- 115  
 Db 80 DLD-----ASRYGSMELKSLISALHGKIQIADVVINHRCADYKDSRGYICIFEG 131  
 QY 116 ----GDLEWNPVNDYTWDFSKVASGKYTANYLDHPNELHAGDSGT--FGGYPDICH-- 168  
 Db 132 GTPDGLDVGPHMICRDBTQPS-----DGTGNL-----DTGADFAAAPDIDHNL 175  
 QY 169 ---DKSNDQYWLWASQBSYAYLSIGIDAWRFYVKGYPVWVKDNLNWWG--GWAUGE 223  
 Db 176 GVVQRELDLWLKSD-----EVGFDARLDFARGYSPEVAKVYIEGTFPVGLAAVE 228  
 QY 224 YWDT-----NVD-----AVLWV-----AYSSGAKVFDFEALYYKMDAEPDNKIPA 263  
 Db 229 LWDNSWYAGDGGKPEYNQDAHQALVDVDRVGGTASAGVYDFTTKIMNTAVEGE--LWR 287  
 QY 264 LVSLALQNGQTVVSDPDPKATFVANHTDIIWNKY-----AYAFILTYEGQPTIFY 315  
 Db 288 LIDQQGKAPGVIGWPAKATFVDNHDGTSTQWMPFPSPDKWQGVAYILTHPGNPCIY 347  
 QY 316 RYDEEWLNKOKLXLIWIHEN---LAGGSTDIVYDNDLIFVRNGVGRKPGILTYINLG 372  
 Db 348 DHFDFGLKQKIALVAVRQNGVTATSKIMLHDAD----- 385  
 QY 373 SSKAGRWVYVPKAGACIHEYTNGLGWVDKYVYSSGWVYLEAPYDPANGQYGVSW 430  
 Db 386 -----AYVAEIDGKVMK-----IG---SRVDVSS-----LIPPGHLLAHNGYAVW 425

RESULT 9

AMC2\_ORYSA STANDARD; PRT; 445 AA.  
 ID AMC2\_ORYSA  
 AC P27941;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amy1ase isozyme C2 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase).  
 GN AMY1.8.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Echaritoidae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 [1]  
 EN SEQUENCE FROM N.A.  
 RP STRAIN=cv. Indica-IR26; TISSUE=Seed;  
 RA Goldman S., Mawal Y., Wu R.;  
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Important for breakdown of endosperm starch during  
 CC germination.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -!- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -!- SUBUNIT: Monomer.  
 CC -!- DEVELOPMENTAL STAGE: Expressed at a high level during germination  
 CC in the aleurones cells under the control of the plant hormone  
 CC gibberellic acid and in the developing grains at a low level.  
 CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X64619; CAA45903.1; --  
 DR PIR; S19990; S19990.  
 DR HSSP; P04063; IAVA.  
 DR Gramene; P27941;  
 DR InterPro; IPR006047; Alpha\_aml1\_cat.  
 DR InterPro; IPR006046; Glyco\_Hydro\_13.  
 DR Pfam; PF00128; alpha-amy1ase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 KW Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 21  
 FT CHAIN 22 445  
 FT ACT\_SITE 202 202  
 FT ACT\_SITE 314 314  
 FT METAL 113 113  
 FT METAL 130 130  
 FT METAL 133 133  
 FT METAL 135 135  
 FT METAL 139 139  
 FT METAL 149 149  
 FT METAL 160 160  
 FT METAL 163 163  
 FT METAL 164 164  
 FT METAL 165 165  
 FT METAL 168 168  
 FT METAL 170 170  
 FT SEQUENCE 445 AA; 49207 MW; DE23701E836ACDA CRC64;  
 Query Match 15.9%; Score 390.5; DB 1; Length 445;  
 Best Local Similarity 27.9%; Pred. No. 8.1e-21;  
 Matches 112; Conservative 61; Mismatches 133; Indels 95; Gaps 17;

QY 12 VIMQAFYWDV--PSGGIWDITROKIPEDYWDAGISAIWIPASKMGCGAYSMGYDPYDF 69  
 Db 24 ILFOGFWESWRQGG--WYNLLMGKVDIVAAAGVTHVWLPFPHSVS---TQYMPGRLY 79  
 QY 70 DLGEYDQGTGTVETFRGSKQELNMINHTAHAYGMKVIADIVINHRAG-----115  
 Db 80 DLD-----ASRYGTSMLKSLISALHGKGIQIADVVINHRCADYKDSRGYCIPEG 131  
 QY 116 ---GDLFWNFVNDYTWDFSKVASKYTYANVLDHFNELHAGDSGT--FGGYDICH-- 168  
 Db 132 GTPDRLDGMGHMTCRDDTQFS-----DGTGNL-----DTGADFAAPDIDHUN 175  
 QY 169 ---DKSNDQYWLWASQESYAYLRSIGIDAWRFDYVKGYAPVWVVDLWNLWVG--GWAVGE 223  
 Db 176 GVQRELTDLWLLKSD-----EVGFDAWLDLPARGYSPEVAKVYIEGTFVGLAVAE 228  
 QY 224 YWDT-----NVD-----AVLW-----AYSSGAKVDFALYKXDEAFDNKNIPA 263  
 Db 229 LWDNSVAYGGDKPYNDQHRQALVDWVDRVGGTASAGVDFDTTKGINVAVEGE-LWR 287  
 QY 264 LVSAALQNGQTVSRDPPKAVTFVANHTDIIWKNIP-----AYAFILTYEGOPTIFY 315  
 Db 288 LIDQGRAPGVIGWPAKAVTFVDNHTDGTSTQOWPSPSKWQGVAYILTHPGNCFIF 347  
 QY 316 RDEEWLNKDKLXNLIWHEN---LAGSTDIYVYNDDELI 353  
 Db 348 DHFFDWGLKQIQAALVAVRQRNGVTATSSLUKIMLHDADAYV 388

## RESULT 10

AMV2\_HORVU STANDARD; PRT; 427 AA.  
 AC P04063;  
 DT 01-NOV-1986 (Rel. 03, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-  
 DE glucan glucanohydrolase) (AMV2-2) (High pI alpha-amylase).  
 GN AMV1.2  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rahmatullah R. J., Huang J. K., Clark K. L., Reese G. R.,  
 RA Chandra G. R., Muthukrishnan S.;  
 RA "Nucleotide and predicted amino acid sequences of two different genes  
 RT for high-pI alpha-amylases from barley.";  
 RL Plant Mol. Biol. 12:119-121(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85131184; PubMed=3871776;  
 RA Rogers J. C.;  
 RA "Two barley alpha-amylase gene families are regulated differently in  
 RT aleurone cells.";  
 RL J. Biol. Chem. 260:3731-3738(1985).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=94254083; PubMed=8196040;  
 RA Kadziola A., Abe J.-I., Svensson B., Haser R.;  
 RA "Crystal and molecular structure of barley alpha-amylase.";  
 RL J. Mol. Biol. 239:104-121(1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH BASI.  
 RC STRAIN=cv. Menuet;  
 RX MEDLINE=98298441; PubMed=9634702;  
 RA Vallee F., Kadziola A., Bourne Y., Juy M., Rodenburg K. W.,  
 RA Svensson B., Haser R.;  
 RA "Barley alpha-amylase bound to its endogenous protein inhibitor BASI:  
 RT crystal structure of the complex at 1.9-A resolution.";  
 RL Structure 6:649-659(1998).

CC -- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -- COFACTOR: Binds 3 calcium ions per subunit.  
 CC -- SUBUNIT: Monomer.  
 CC -- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally  
 CC regulated. Germinating embryos produce the hormone gibberellic  
 CC acid, which within 10 hours stimulates the aleurone cells covering  
 CC the endosperm of the seed to produce alpha-amylase. The enzyme  
 CC then degrades the starch within the endosperm for use by the  
 CC developing plant embryo.  
 CC -- INDUCTION: Type B isozyme mRNA is undetectable in unstimulated  
 CC cells and increases a hundred-fold after stimulation with  
 CC gibberellic acid.  
 CC -- MISCELLANEOUS: There are at least 4 types of alpha-amylase in  
 CC barley.  
 CC -- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X15226; CA33298.1; -;  
 CC EMBL; K02637; AAA98790.1; -;  
 CC PIR; A31960; ALBHB.  
 CC PDB; 1AMY; 13-MAY-95.  
 CC PDB; 1AVA; 16-MAR-99.  
 CC PDB; 1BG9; 15-JUN-99.  
 CC InterPro; IPR006589; Alp\_amyl\_cat.sub.  
 CC InterPro; IPR006047; Alpha\_amyl\_cat.  
 CC InterPro; IPR006046; Glyco\_Hydro\_13.  
 CC Pfam; PF00128; alpha-amylase; 1.  
 CC PRINTS; PR00110; ALPHAAMYLASE.  
 CC SMART; SM00642; Amy; 1.  
 CC Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
 CC Calcium-binding; signal; Multigene family; 3D-structure.  
 CC SIGNAL 1 24 ALPHA-AMYLASE TYPE B ISOZYME.  
 FT CHAIN 25 427  
 FT ACT\_SITE 203 203  
 FT ACT\_SITE 228 228  
 FT ACT\_SITE 313 313  
 FT METAL 115 115  
 FT METAL 132 132  
 FT METAL 135 135  
 FT METAL 137 137  
 FT METAL 141 141  
 FT METAL 151 151  
 FT METAL 162 162  
 FT METAL 165 165  
 FT METAL 166 166  
 FT METAL 167 167  
 FT METAL 170 170  
 FT METAL 172 172  
 FT CONFLICT 134 134  
 FT CONFLICT 195 197  
 FT CONFLICT 425 425  
 FT STRAND 27 29  
 FT TURN 33 34  
 FT HELIX 35 37  
 FT TURN 39 40  
 FT HELIX 42 47  
 FT TURN 48 49  
 FT HELIX 50 56  
 FT TURN 57 57  
 FT STRAND 60 63  
 FT TURN 69 69  
 FT TURN 72 73  
 FT STRAND 77 77  
 FT TURN 80 81  
 FT TURN 84 85



PT TURN 87 88  
 FT TURN 91 104  
 FT TURN 105 105  
 FT STRAND 107 112  
 FT STRAND 116 116  
 FT STRAND 121 122  
 FT TURN 124 125  
 FT STRAND 128 130  
 FT TURN 139 140  
 FT HELIX 144 146  
 FT STRAND 147 147  
 FT TURN 149 150  
 FT TURN 152 154  
 FT STRAND 165 165  
 FT TURN 168 169  
 FT STRAND 172 173  
 FT TURN 175 176  
 FT HELIX 178 193  
 FT TURN 194 195  
 FT STRAND 199 202  
 FT TURN 203 204  
 FT HELIX 205 207  
 FT TURN 210 220  
 FT STRAND 224 227  
 FT STRAND 235 235  
 FT TURN 237 238  
 FT STRAND 241 241  
 FT HELIX 245 260  
 FT TURN 261 262  
 FT STRAND 265 268  
 FT TURN 270 279  
 FT HELIX 280 282  
 FT TURN 284 287  
 FT TURN 290 291  
 FT HELIX 297 299  
 FT TURN 300 300  
 FT TURN 302 304  
 FT STRAND 305 308  
 FT TURN 312 314  
 FT TURN 316 318  
 FT HELIX 325 327  
 FT HELIX 328 337  
 FT STRAND 341 345  
 FT HELIX 346 350  
 FT TURN 351 351  
 FT HELIX 355 367  
 FT TURN 368 369  
 FT TURN 372 373  
 FT STRAND 376 382  
 FT TURN 383 384  
 FT STRAND 385 390  
 FT TURN 391 393  
 FT STRAND 394 398  
 FT HELIX 405 407  
 FT STRAND 412 418  
 FT TURN 419 420  
 FT STRAND 421 427  
 SQ SEQUENCE 427 AA; 47355 MW; 957C0B16621BF748 CRC64;

Query Match 14.9%; Score 366.5; DB 1; Length 427;  
 Best Local Similarity 26.8%; Pred. No. 4.1e-19;  
 Matches 114; Conservative 60; Mismatches 132; Indels 123; Gaps 19;

QY 5 SELEKGVIMQAFYWDV-PSGGIWDITRQKIPWYDAGISAIWIPPASKMGGAYSNGY 63  
 DB 19 ASLASQVLFPQGNWESKXNGWYFLGKYDDIAAGITHVLPFPASQSWA---EGY 75  
 QY 64 DPYDFDLGDKYDQGVTRFSGKQELVNMINTAHYGNKVIADIVINHRA----- 114  
 DB 76 MPCRLYDL-----ASKYGNKAQLKSLIGALHGKGVKATADIVINHTAEHKGGRGI 127  
 QY 115 -----LEWNPV---NDYTWTFPSKVASKGYTANYLDFPNELHAGDSGT-FG 161

DB 128 YCIFEGGTPDARLDWGPMMICRRDRPYAD-----GTGN-----PTGADFG 168  
 QY 162 GYPDICH-----DKSWDQYVLMWASQESYAAVLRISIGIDAMRFYVYKGYAPVWVKDMLN-W 215  
 DB 169 AAPDIDHLNLRVQKELVWLNWKAD-----IGPDGWRFDPAKGYSDADVAKIYIDRS 220  
 QY 216 WGVNAVGEYW-----DTNVDVNLWAYSSGAK-----VDFPALYKMDFAF 256  
 DB 221 EFSFAVAETWTSLAYGGDKPNLNQDHRQLVNMVVKVGGKGPATTFDFTTKGILNAV 280  
 QY 257 DNK-----NIPALYSALQNGQTVVSRDPFKAFTFVANHTDIIWNKYP----- 299  
 DB 281 EGEMLNRLGTDCGKAPGMIGW-----PAKAVTFVDNEDTGSTQHMWFFSDRVMQ 330  
 QY 300 AYAFILTYEGQTFIYROYEVLNWKDKLNLI-----WIHE-----NLGSGSTIVVYDN 349  
 DB 331 GYAYILTHEGTCIFDHFDFDGLKKEIDRLVSRVTRHGHIHNEKSKLOITIEADADLYLAEI 390  
 QY 350 DELIFVRNG 358  
 DB 391 DGKVIKLG 399

RESULT 11  
 AMYL\_HORVU  
 ID AMYL\_HORVU STANDARD; PRT: 438 AA.  
 AC P00633;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type A isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase) (AMYL) (Low pI alpha-amylase).  
 GN AMYL1.1.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.  
 OC NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Himalaya;  
 RX MEDLINE=83238423; PubMed=6190808;  
 RA Rogers J.C., Millman C.;  
 RT "Isolation and sequence analysis of a barley alpha-amylase cDNA clone.";  
 RL J. Biol. Chem. 258:8169-8174(1983).  
 CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.  
 CC -|- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -|- SUBUNIT: Monomer (By similarity).  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally regulated. Germinating embryos produce the hormone gibberellic acid, which within 10 hours stimulates the aleurone cells covering the endosperm of the seed to produce alpha-amylase. The enzyme then degrades the starch within the endosperm for use by the developing plant embryo.  
 CC -|- MISCELLANEOUS: There are at least 4 types of alpha-amylase in barley.  
 CC -|- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J01236; AAA32929.1; -.  
 CC PIR: A00846; ALEH.  
 CC HSSP: P04063; IAVA.  
 CC InterPro: IPR006589; Alp\_amyl\_cat\_sub.

DR InterPro; IPR006047; Alpha amyl cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SMO0642; Amy; 1.  
 KW Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
 KW Calcium-binding; Signal; Multigene family.  
 FT SIGNAL 1 24  
 FT CHAIN 25 438 ALPHA-AMYLASE TYPE A ISOZYME.  
 FT ACT\_SITE 204 204 BY SIMILARITY.  
 FT ACT\_SITE 229 229 BY SIMILARITY.  
 FT ACT\_SITE 315 315 BY SIMILARITY.  
 FT METAL 116 116 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 133 133 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 136 136 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 138 138 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 142 142 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 152 152 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 163 163 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 166 166 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 167 167 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 168 168 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 171 171 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 173 173 CALCIUM 1 AND 3 (BY SIMILARITY).  
 FT SEQUENCE 438 AA; 47796 MW; 2393FDAC51B80F51 CRC64;  
 Query Match 14.9%; Score 365.5; DB 1; Length 438;  
 Best Local Similarity 27.2%; Pred. No. 4.9e-19;  
 Matches 109; Conservative 64; Mismatches 132; Indels 95; Gaps 18;  
 QY 12 VIMQAFWD--VPSGGIWDITROKIPENYDAGISAIWTPASKMGGAYSGYDPDF 69  
 DB 27 VLFQGNWESKQSGG-WYNNMMKVDVIAAGVTHVLPSPSHSVS---NEGYPGRLY 82  
 QY 70 DLGBYDQKGTVEFRGSKQELVNMNINAHAYGMKVIADIVINHR----- 113  
 DB 83 DID-----ASYGNAELKSLIGALHGKGVQAIADIVINHRCAADYKDSRGICVIEG 134  
 QY 114 --AGDLEWNPVNDYTWTPSKVASKYTANKYLDHFHNLHAGDSGT-FGYPDICH-- 168  
 DB 135 GTSQRLDWGSHM---ICRDTKYSDG--TANL-----DTGADFAAAPDIDHLN 178  
 QY 169 ---DKSWQVWLWASQBSYAYLRSIGIDAWRFDYKGYAPVWVMDLNNWVG-GNAVGEY 224  
 DB 179 DRVQRELKWLWLKSD-----LGFDAWRLDFARGYSPENAKYVIDGTSPSLAVAEV 230  
 QY 225 WDT-----NVDA-----VLNW-----AYSSGAKVDFPALYVMDAEPDNKNIPAL 264  
 DB 231 WDNMATGDKGNPDODAHQNLVNWVDRKVGGAASAGWVDFTTKGLNAAVEGE-LMRL 289  
 QY 265 VSALQNGQTVVSRDPFKAVTFVANHDTDIWNKYP-----AVAFILTYEGQPTIFVR 316  
 DB 290 IDPGKAPGVGWGPKAAATFVNDHGTSTQAMWPPSKWQGYAYILTHPGICIFVD 349  
 QY 317 DYEWLKNDKLNLIWHEN---LAGGSTDIVYDNDLI 353  
 DB 350 HFFWNGFQDQIALVAIRKNGITATSAKILMHGDAYV 389  
 RESULT 12  
 ID AMY3 HORVU  
 AC P04747; STANDARD; PRT; 368 AA.  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-  
 DE glucan glucanohydrolase) (Clone PHV19) (Fragment).  
 GN AMYL3.  
 OS Hordeum vulgare (Barley).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
 Triticeae; Hordeum.  
 NCBI\_TaxID=4513;  
 (1)  
 RP SEQUENCE FROM N.A.  
 RA Chandler P.M., Zwar J.A., Jacobsen J.V., Higgins T.J.V., Inglis A.S.;  
 "The effects of gibberellic acid and abscisic acid on alpha-amylase  
 mRNA levels in barley aleurone layers studies using an alpha amylase  
 cDNA clone."; Plant Mol. Biol. 3:407-418(1984).  
 RL -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 linkages in oligosaccharides and polysaccharides.  
 CC -1- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally  
 regulated. Germinating embryos produce the hormone gibberellic  
 acid, which within 10 hours stimulates the aleurone cells covering  
 the endosperm of the seed to produce alpha-amylase. The enzyme  
 then degrades the starch within the endosperm for use by the  
 developing plant embryo.  
 CC -1- MISCELLANEOUS: There are at least 4 types of alpha-amylase in  
 barley.  
 CC -1- MISCELLANEOUS: Type B isozyme mRNA is undetectable in unstimulated  
 cells and increases a hundred-fold after stimulation with  
 gibberellic acid.  
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 EMBL; K02638; AAA32933.1; -.  
 HSBL; P04063; IAVA.  
 InterPro; IPR006589; Alp amyl cat sub.  
 InterPro; IPR006047; Alpha amyl cat.  
 InterPro; IPR006046; Glyco\_hydro\_13.  
 Pfam; PF00128; alpha-amylase; 1.  
 PRINTS; PR00110; ALPHAAMYLASE.  
 SMART; SMO0642; Amy; 1.  
 KW Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
 KW Calcium-binding; Signal; Multigene family.  
 FT SIGNAL 1 24  
 FT CHAIN 25 438 ALPHA-AMYLASE TYPE B ISOZYME.  
 FT ACT\_SITE 203 203 BY SIMILARITY.  
 FT ACT\_SITE 228 228 BY SIMILARITY.  
 FT ACT\_SITE 313 313 BY SIMILARITY.  
 FT METAL 115 115 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 132 132 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 135 135 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 137 137 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 141 141 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 151 151 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 162 162 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 165 165 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 166 166 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 167 167 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 170 170 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 172 172 CALCIUM 1 AND 3 (BY SIMILARITY).  
 FT NON\_TER 368 368  
 FT SEQUENCE 368 AA; 40787 MW; A237EF55793BA93B CRC64;  
 Query Match 14.8%; Score 363.5; DB 1; Length 368;  
 Best Local Similarity 27.0%; Pred. No. 5.6e-19;  
 Matches 106; Conservative 55; Mismatches 118; Indels 113; Gaps 16;

QY 5 SELEKGVIMQAFYWDV-PSGGIWDITROKIPWYDAGISAIWIPASKMGGAYSMGY 63  
Db 19 ASLASQVLFQGFNWSKNGWYNFLNGKVDYDIAAGITHWLPFASQSV-EGQY 75  
QY 64 DYVDFDLGEYQKGTVEFRFSKQELVNMINTAHAYGMKVIADIVINHRA- 115  
Db 76 MPGRLYDLD-ASKYGNKQKSLIGALHGKGVKALADIVINHRTAEHKGDRGI 127  
QY 116 -GOLWNPFFV-NDVTWTFDSKVASKYNTANYLDHFNELHAGDSGT-RG 161  
Db 128 YCIFEFGVTPDARLDWCPHMCDDRPYAD-GTGN-PDTGADFG 168  
QY 162 GYDPDICH-DKSWDQVWLWASQESYAALRSIGIDAWRFYVKGYPVVKDWLN-W 215  
Db 169 AAPDIDHMLRVLQKELAELWNLKAD-IGFDGWRDFFAKGSADVAKIYIDRS 220  
QY 216 WCGWAVEYV-DTNVDVAVLNWAYSSGAK-VDFPALLYKMDKDEAF 256  
Db 221 ERSFAVAEWTSLAYGGCKPNLQOQRQELVNVWVDKVGKGPATTFDTTKGILNAV 280  
QY 257 DNK-NIPALVSALQNGQTVVSRDPFKAVTFVANHTDIIWVKYP- 299  
Db 281 EGELWRLRGTDGKAPGMIGW-PAKAVTFVDNHDGTSGTQHMPPFSDRVWQ 330  
QY 300 AVAFILTVGQPTIFRYDYEEWLNKDKLXLI 331  
Db 331 GYAILTHPGTFCIFYDHFDDWGLKEBIDRLV 362

RESULT 13  
ID AM3B ORYSA  
AC P27937; STANDARD; PRT; 438 AA.  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Alpha-amylase isozyme 3B precursor (RC 3.2.1.1) (1,4-alpha-D-glucan  
DE glucanohydrolase).  
GN AMV1.6 OR AMV3B.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]\_\_SEQUENCE FROM N.A.  
RC STRAIN=sv. Japonica M202; TISSUE=Etisolated leaf;  
RX MEDLINE=91329692; PubMed=1714318;  
RA Sutliff T.D., Huang N., Litts J.C., Rodriguez R.L.;  
RT "Characterization of an alpha-amylase multigene cluster in rice."  
RL Plant Mol. Biol. 16:579-591(1991).  
RN [2]\_\_SEQUENCE FROM N.A.  
RA Sutliff T.D., Huang N., Rodriguez R.L.;  
RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Important for breakdown of endosperm starch during  
CC germination.  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
CC linkages in oligosaccharides and polysaccharides.  
CC -!- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
CC -!- SUBUNIT: Monomer.  
CC -!- TISSUE SPECIFICITY: Germinating seeds.  
CC -!- DEVELOPMENTAL STAGE: Expressed at a high level during germination  
CC in the aleurone cells under the control of the plant hormone  
CC gibberellic acid and in the developing grains at a low level.  
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
EMBL; X56337; CAA39777.1; -  
EMBL; M24941; AAA33897.1; -  
PIR; S14957; S14957.  
HSP; P04063; IAVA.  
Gramene; P27937; -  
InterPro; IPR006589; Alp\_aml\_cat\_sub.  
InterPro; IPR006047; Alpha\_aml\_cat.  
InterPro; IPR006046; Glyco\_hydro\_13.  
Pfam; PF00128; alpha-amylase; 1.  
PRINTS; PR00110; ALPHAAMYLASE.  
SMART; SM00642; Amy; 1.  
Carbohydrate metabolism; Hydrolase; Calcium-binding;  
Signal; Multigene family.  
FT SIGNAL 1 26  
FT CHAIN 27 438  
FT ACT\_SITE 205 205  
FT ACT\_SITE 313 313  
FT METAL 117 117  
FT METAL 134 134  
FT METAL 137 137  
FT METAL 139 139  
FT METAL 143 143  
FT METAL 153 153  
FT METAL 164 164  
FT METAL 167 167  
FT METAL 168 168  
FT METAL 169 169  
FT METAL 172 172  
FT METAL 174 174  
FT SEQUENCE 438 AA; 48591 MW; B9DE0B5ABC63F9C CRC64;  
SQ  
Query Match 14.5%; Score 357.5; DB 1; Length 438;  
Best Local Similarity 27.2%; Pred. No. 1.8e-18;  
Matches 105; Conservative 53; Mismatches 123; Indels 105; Gaps 17;  
QY 5 SELEKGVIMQAFYWD-VPSGGIWDITROKIPWYDAGISAIWIPASKMGGAYSMG 62  
Db 21 SHLAQVLFQGFNWSKQGG-WYNFLHGHVDDIAATGVTHVWLPFPHSVA-PQG 76  
QY 63 YDPYDFDLGEYQKGTVEFRFSKQELVNMINTAHAYGMKVIADIVINHRA- 114  
Db 77 YMGRLYDLD-ASKYGTGAELRSLAAFSKGIKCVADIVINHRCADYKDSRG 128  
QY 115 -GGD-LEWNPVNDYTWTFDSKVASKYNTANYLDHFNELHAGDSGTGGYVP 164  
Db 129 IYCFEGGTGPDRLDWP--DMICSDDTQYNGRG-HRDTGADFGAAP 173  
QY 165 DICHKSDQY-WL-WASQESYAALRSIGIDAWRFYVKGYPVVKDWLN-NWGG 218  
Db 174 DIDHLNTRVQTELSDLNLWLKSD-VGFDGWRDLDFAGYSAVAKTYVDNTDPS 225  
QY 219 WAVEYV-DTNVDVAVLNWAYSSG-AKVDFPALLYKMDKDEAF- 257  
Db 226 FVVAEITWSNMRYDNGEPEWNGDGRQELVNVWVAQVGGPASAFDTTKGELQAAVQGEW 285  
QY 258 -NKNIPALVSALQNGQTVVSRDPFKAVTFVANHTDIIWVKYP-AYAPI 304  
Db 286 RMDKNGKAPGMIGWL-PEKAVTFIDNHDGTSGTQNSWPPFSDKVMQGYAI 335  
QY 305 LTYEGQPTIFRYDYEEWLNKDKLXLI 330  
Db 336 LTHPGVPCIFYDHFVWNLKQEISTL 361  
RESULT 14  
ID AM3C ORYSA  
AC P27939; STANDARD; PRT; 437 AA.

DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Alpha-amylase isozyme 3C precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase).  
 GN AMY1.7 OR AMY3B.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Japonica M202; TISSUE=etiolated leaf;  
 RX MEDLINE=91328692; PubMed=1714318;  
 RA Sutliff T.D., Huang N., Litts J.C., Rodriguez R.L.;  
 RT "Characterization of an alpha-amylase multigene cluster in rice.";  
 RL Plant Mol. Biol. 16:579-591(1991).  
 CC -!- FUNCTION: Important for breakdown of endosperm starch during  
 CC germination.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -!- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -!- SUBUNIT: Monomer.  
 CC -!- TISSUE SPECIFICITY: Germinating seeds.  
 CC -!- DEVELOPMENTAL STAGE: Expressed at a high level during germination  
 CC in the aleurones calls under the control of the plant hormone  
 CC gibberellic acid and in the developing grains at a low level.  
 CC -!- SIMILARITY: Belongs to family 13 of Glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X56338; CAA39778.1; -  
 DR PIR; S14956; S14956.  
 DR HSSP; P04063; 1A.VA.  
 DR Gramene; P27339; -  
 DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
 DR InterPro; IPR006047; Alp\_amy1\_cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Aamy; 1.  
 DR Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 26  
 FT CHAIN 27 437  
 FT ACT SITE 205 205  
 FT ACT SITE 313 313  
 FT ACT SITE 117 117  
 FT METAL 134 134  
 FT METAL 137 137  
 FT METAL 139 139  
 FT METAL 143 143  
 FT METAL 153 153  
 FT METAL 164 164  
 FT METAL 167 167  
 FT METAL 168 168  
 FT METAL 169 169  
 FT METAL 172 172  
 FT METAL 174 174  
 FT METAL 174 174  
 SQ SEQUENCE 437 AA; 48637 MW; BD304250B40C7A8B CRC64;  
 Query Match 14.4%; Score 353; DB 1; Length 437;  
 Best Local Similarity 26.4%; Pred. No. 3.9e-18;

Matches 108; Conservative 55; Mismatches 138; Indels 108; Gaps 18;  
 QY 5 SELEKGVIMQAFYNDV--PSGGIWWDTIRQKIPWYDAGISAIWIPASKCMGAYSMG 62  
 DB 21 SHLAQAQVLFQGFNWSNKKQG-WYNFLHSHVDYIAATGTVHWLPPPSHVA---PQG 76  
 QY 63 YDPYDFDLCEYDQKGTVEFRGSKGLVNMINTAHAYGMKVIADIVINHRA----- 114  
 DB 77 YMPGRLYDLD-----ASKYGTGAELSLIAAFHSHKSIKCAVIVINHRCADYKDSRG 128  
 QY 115 -----GGD-----LEWNPFFVNDYTWDFSKVAGSKYTANYLDFHFNELHAGDSGTFGGYP 164  
 DB 129 IYCFEGGTPDSRLDWGP---DMICSDDTQYNSGRG-----HRTGADFGAAP 173  
 QY 165 DICHKDSWDQY---WL-WASQESYAAVLRISIGIDAWREDYVYKGYAPVWVQDWL-NMGG 218  
 DB 174 DIDHLNTRVQTELSDLNLKSD-----VGFDGWRLLDPAKGYSATVAKTYVDNTDPS 225  
 QY 219 WAYGEYH-----DTNVDVAVNWAYSSG--AKVDFALYKXDEAFD----- 257  
 DB 226 FVVAETSMNRYDNGEPPSNQDGRQELVNAQAVGGPASFDTTKGELQAAVQGSWM 285  
 QY 258 -----NKNIPALVSALQNGQTVVSRDPFRAVTFVANHDTDIINNKYP-----AYAFI 304  
 DB 286 RMKDGNGKAPGMIGWL-----PEKAVTFIDNHDGTGTCNSWPPFSDKVMQRYAVI 335  
 QY 305 LTVEGOPTFYRDEYEWLNKDKLNIWHENLA---GGSTDIVYVYDND 350  
 DB 336 LTHFGVPCFYDHFVDFWNLIKQEIETLAAVRSRNGIHGSKNLILAAADGD 384  
 RESULT 15  
 ID AMY6 HORVU STANDARD; PRT; 429 AA.  
 AC P04750;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-  
 DE glucan glucanohydrolase) (Clones GRAMy56 and 963).  
 GN AMY1.6.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A. (CLONE GRAMy56).  
 RA Rahmatullah R.J., Huang J.-K., Clark K.L., Reeck G.R.,  
 RA Chandra G.R., Muthukrishnan S.;  
 RT "Nucleotide and predicted amino acid sequences of two different genes  
 RT for high-pi alpha-amylases from barley.";  
 RL Plant Mol. Biol. 12:119-121(1989).  
 RN [2]  
 RP SEQUENCE OF 380-429 FROM N.A. (CLONE 963).  
 RX MEDLINE=85159405; PubMed=6335720;  
 RA Huang J.-K., Swegle M., Dandekar A.M., Muthukrishnan S.;  
 RT "Expression and regulation of alpha-amylase gene family in barley  
 RT aleurones.";  
 RL J. Mol. Appl. Genet. 2:579-588(1984).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -!- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally  
 CC regulated. Germinating embryos produce the hormone gibberellic  
 CC acid, which within 10 hours stimulates the aleurone cells covering  
 CC the endosperm of the seed to produce alpha-amylase. The enzyme  
 CC then degrades the starch within the endosperm for use by the  
 CC developing plant embryo.  
 CC -!- MISCELLANEOUS: There are at least 4 types of alpha-amylase in  
 CC barley.  
 CC -!- MISCELLANEOUS: Type B isozyme mRNA is undetectable in unstimulated

CC cells and increases a hundred-fold after stimulation with  
CC gibberellic acid.  
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X15227; CAA33299.1; -;  
CC EMBL; K02636; AAB32932.1; -;  
CC PIR; J04067; J04067.  
CC HSP; P04063; IAVA.  
CC InterPro; IPR006589; Alp.amyl cat sub.  
CC InterPro; IPR006047; Alpha.amyl cat.  
CC InterPro; IPR006046; Glyco\_hydro\_13.  
CC Pfam; PF00128; alpha-amylase; 1.  
CC PRINTS; PR00110; ALPHAAMYLASE.  
CC SMART; SM00642; Amy; 1.  
KW Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
KW Calcium-binding; Signal; Multigene family.  
FT SIGNAL 1 24  
FT CHAIN 25 429 ALPHA-AMYLASE TYPE B ISOZYME.  
FT ACT\_SITE 205 205 BY SIMILARITY.  
FT ACT\_SITE 315 315 BY SIMILARITY.  
FT METAL 115 115 CALCIUM 1 (BY SIMILARITY).  
FT METAL 132 132 CALCIUM 2 (BY SIMILARITY).  
FT METAL 135 135 CALCIUM 2 (BY SIMILARITY).  
FT METAL 137 137 CALCIUM 2 (BY SIMILARITY).  
FT METAL 141 141 CALCIUM 2 (BY SIMILARITY).  
FT METAL 151 151 CALCIUM 3 (BY SIMILARITY).  
FT METAL 167 167 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY  
FT SIMILARITY).  
FT METAL 168 168 CALCIUM 1 (BY SIMILARITY).  
FT METAL 169 169 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY  
FT SIMILARITY).  
FT METAL 172 172 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY  
FT SIMILARITY).  
FT METAL 174 174 CALCIUM 1 AND 3 (BY SIMILARITY).  
FT SEQUENCE 429 AA; 47937 MW; 1C924CA6319D5262 CRC64;  
Query Match 14.3%; Score 351.5; DB 1; Length 429;  
Best Local Similarity 27.1%; Pred. No. 4.8e-18;  
Matches 106; Conservative 54; Mismatches 122; Indels 109; Gaps 17;  
QY 7 LEKGVIVQAFYWDV-PSGGIWDITROKIPEDWDGASAIWIPPSKMGGMGYDYP 65  
DB 21 LASGVLFQGENWESKNGWYFNLMGKVDDIAAGVTHVWLPFASQSA---EQGYMP 77  
QY 66 YDFFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVNHRA----- 114  
DB 78 GELYLD-----ASKYGNKAQLKSLIGALHGKAVKAIADIVINHRTAERKDGRIYC 129  
QY 115 ---GSD---LEWNPV---NDYTWTFDSKVASKYTANYLDHFNELHAGDSGTFGYYP 164  
DB 130 IFEGGTPDARLDWGHMCRDRPYPD-----GTGN---RPTRTA---DFGAAP 173  
QY 165 DICH-----DKSDQYWLWASQESYAAYLRISIGIDARRFDYVGYAPWVKDWLN-WWGG 218  
DB 174 DIDHLPVQKELVEWLNLRITDD-----GFGWRFDFAKGSADVAKIYYDRSEPS 225  
QY 219 WAVEGYW-----DINVDVNLWAYSSG-----AKVDFPALYYKMDFAFDNK 259  
DB 226 FAVAEIWTSLAYGGDGKFNQDHRQELVNMVNVKVGSGPATTFDTTKGILNVAVEGE 285  
QY 260 -----NIPALVSALONGQTVVSRDPFKAVTFVANHDTIWNKYP-----AYA 302  
DB 286 LWRLEGTGKAPGMIGW-----PAKAVTFVDNHDGTGQHMWPFPSDRWMOGYA 335  
QY 303 FILTYEGQPTFYRDYEEWLNKDKLNLIWI 333

Db 336 YILTHPGNPCIYDFDHFDFWGLKEEDRLVSI 366  
Search completed: June 29, 2004, 09:34:44  
Job time : 19 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:31:27 ; Search time 45 seconds  
(without alignments)  
3057.021 Million cell updates/sec

Title: US-10-081-739A-2  
Perfect score: 2459  
Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVNSVSGVG 436

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mmc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2311	94.0	432	14 Q8JZK3	Q8jzk3 uncultured
2	2298	93.5	461	1 Q8NKR5	Q8nkr5 thermococcus
3	2279	92.7	461	1 Q8NKR4	Q8nkr4 thermococcus
4	2248	91.4	461	1 Q33476	Q33476 pyrococcus
5	2221	90.3	457	1 Q93647	Q93647 thermococcus
6	2199.5	89.4	469	1 Q50200	Q50200 thermococcus
7	2153	87.6	460	1 Q84522	Q84522 pyrococcus
8	2153	87.6	473	17 Q8U319	Q8u319 pyrococcus
9	2152	87.5	460	1 Q9P910	Q9p910 pyrococcus
10	513	20.9	519	2 Q9Q7R8	Q9q7r8 cytophaga s
11	511.5	20.8	507	16 Q87HG6	Q87hg6 vibrio paka
12	492	20.0	492	16 Q8YUZ1	Q8yuz1 anabaena sp
13	487.5	19.8	494	16 Q8Z5S5	Q8z5s5 salmonella
14	486	19.8	501	2 Q3I148	Q3i148 bacillus sp
15	472.5	19.2	516	2 Q82839	Q82839 bacillus sp
16	469.5	19.1	549	2 Q31193	Q31193 bacillus st

17	468.5	19.1	549	2	Q9KXV6	Q9kxv6 bacillus st
18	467.5	19.0	531	16	Q8IAS4	Q8ias4 bacillus ce
19	464.5	18.9	521	2	P71034	P71034 bacillus sp
20	464.5	18.9	533	2	Q9AQ54	Q9aq54 bacillus me
21	456	18.5	613	2	Q59222	Q59222 bacillus sp
22	455.5	18.5	513	16	Q81YJ4	Q81yj4 bacillus an
23	452.5	18.4	495	16	Q8XBB6	Q8xbb6 escherichia
24	451.5	18.4	495	16	Q8JR40	Q8jr40 shigella fl
25	451.5	18.4	506	16	Q8U916	Q8u916 agrobacteri
26	448.5	18.2	495	16	Q8FGL8	Q8fgl8 escherichia
27	446.5	18.2	495	16	Q7UAB0	Q7uab0 shigella fl
28	438	17.8	491	16	Q9CG59	Q9cg59 lactococcus
29	435	17.7	481	16	Q89YPI	Q89yp1 bacteroides
30	420	17.1	493	2	Q03657	Q03657 bacillus ci
31	412.5	16.8	420	10	Q92P43	Q92p43 phaseolus v
32	406	16.5	421	10	Q7X9T1	Q7x9t1 phaseolus a
33	398	16.2	423	10	Q42678	Q42678 cuscuta ref
34	392	15.9	424	10	Q8LP27	Q8lp27 pharbitis n
35	385.5	15.7	488	16	Q8E696	Q8e696 streptococ
36	385.5	15.7	488	16	Q8E0M2	Q8e0m2 streptococ
37	379.5	15.4	484	2	O50583	O50583 streptococ
38	376.5	15.3	421	10	Q42504	Q42504 hordeum vul
39	373.5	15.2	437	10	O44965	O44965 hordeum vul
40	373	15.2	484	16	Q97Q49	Q97q49 streptococ
41	372	15.1	484	16	Q8DPC8	Q8dpc8 streptococ
42	371	15.1	403	10	Q7Y1C3	Q7y1c3 eleusine co
43	365.5	14.9	407	10	O41442	O41442 solanum tub
44	363.5	14.8	504	1	Q60224	Q60224 natronococ
45	362	14.7	416	10	Q8LJQ6	Q8ljq6 musa acumin

ALIGNMENTS

RESULT 1  
Q8JZK3 PRELIMINARY; PRT; 432 AA.  
ID Q8JZK3  
AC Q8JZK3;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-amylase.  
GN B5063.  
OS uncultured organism.  
OC unclassified; environmental samples.  
OX NCBI\_TaxID=155900;  
RN [1]  
RP MEDLINE-22113025; PubMed-11994309;  
RX Richardson T.H., Fan X., Frey G., Callen W., Cabell M., Lam D.,  
RA Macomber J., Short J.M., Robertson D.E., Miller C.;  
RT "A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE.";  
J. Biol. Chem. 277:26501-26507(2002).  
RL EMBL: AF504064; AAM48114.1; -;  
DR GO: GO:0004556; F:alpha-amylase activity; IEA.  
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro: IPR006047; Alpha\_amyl\_cat.  
DR InterPro: IPR006589; Alp\_amyl\_cat\_sub.  
DR Pfam: PF00128; Glyco\_hydro\_13.  
DR Pfam: PF00128; alpha-amylase; 1.  
DR PRINTS: PR00110; ALPHAAMYLASE.  
DR SMART: SM00642; Amy; 1.  
SQ SEQUENCE 432 AA; 49225 MW; EC5A87826D31D2D CRC64;

Query Match 94.0%; Score 2311; DB 14; Length 432;  
Best Local Similarity 94.7%; Pred. No. 3e-148;  
Matches 407; Conservative 10; Mismatches 13; Indels 0; Gaps 0;  
QY 7 LEKGGVIMQAFYWDVPSGGIWDWDTIRQKIPEDWDAGISAIWIPDPASKMGSGYSGYDPY 66  
DB 3 LEEGGLIMQAFYWDVPMGGIWDWDTIAQKIPDWASAGISAIWIPDPASKMGSGYSGYDPY 62

QY 67 DPFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWNPFFVND 126  
 Db 63 DPFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWNPFFVND 122  
 QY 127 YTWDFSKVASGKYTANYLDLDFHFNELHAGDSGTGGYDPDI CHDKSWDQYWLWASQ 186  
 Db 123 YTWDFSKVASGKYTANYLDLDFHFNELHAGDSGTGGYDPDI CHDKSWDQYWLWASQ 182  
 QY 187 YLRSIGIDAWRFDYVKGYPVYKDWLWNGWGWAVEYDNTVDVAVLNWYSSGAKVDF 246  
 Db 183 YLRSIGIDAWRFDYVKGYPVYKDWLWNGWGWAVEYDNTVDVAVLNWYSSGAKVDF 242  
 QY 247 ALYKQDEAFDNKNI PALYSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPAYAFILT 306  
 Db 243 ALYKQDEAFDNKNI PALVDALRYGTVVSRDPFKAVTFVANHTDIIWNKYPAYAFILT 302  
 QY 307 YEQOPTIFVRDYEEWLNKDLKNIWIHENLAGGSTDIVYDNDDELIFVRNGYGD 366  
 Db 303 YEQOPTIFVRDYEEWLNKDLKNIWIHENLAGGSTDIVYDNDDELIFVRNGYGD 362  
 QY 367 TYINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPANGQY 426  
 Db 363 TYINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPANGQY 422  
 QY 427 YSVWSYCGVG 436  
 Db 423 YSVWSYCGVG 432

## RESULT 2

Q8NKR5 Q8NKR5 PRELIMINARY; PRT; 461 AA.  
 ID Q8NKR5  
 AC Q8NKR5  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Alpha-amylase precursor.  
 GN BD5064.  
 OS Thermococcus sp. GUSL5.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Thermococcus.  
 OX NCBI\_TaxID=195100;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GUSL5;  
 RX MEDLINE=22113025; PubMed=11994309;  
 RA Richardson T.H., Tan X., Frey G., Callen W., Cabell M., Lam D.,  
 RA Macomber J., Short J.M., Robertson D.E., Miller C.;  
 RT "A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
 RT AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE.";  
 RL J. Biol. Chem. 277:26501-26507(2002).  
 DR EMBL; AF504062; AA48112.1; -;  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Amy; 1.  
 KW Signal.  
 FT SIGNAL  
 SQ SEQUENCE 1 26 POTENTIAL.  
 Query Match 93.5%; Score 2298; DB 1; Length 461;  
 Best Local Similarity 93.1%; Pred. No. 2.5e-147;  
 Matches 405; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 2 AKYSELEKGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGWGAYSM 61  
 Db 27 AKYSELEKGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGWGAYSM 86  
 QY 62 GYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWNP 121

Db 87 GYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWNP 146  
 QY 122 PFVNDYTWDFSKVASGKYTANYLDLDFHFNELHAGDSGTGGYDPDI CHDKSWDQYWLWASQ 181  
 Db 147 PFVNDYTWDFSKVASGKYTANYLDLDFHFNELHAGDSGTGGYDPDI CHDKSWDQYWLWASQ 206  
 QY 182 ESAAALYRSIGIDAWRFDYVKGYPVYKDWLWNGWGWAVEYDNTVDVAVLNWYSSGA 241  
 Db 207 ESAAALYRSIGIDAWRFDYVKGYPVYKDWLWNGWGWAVEYDNTVDVAVLNWYSSGA 266  
 QY 242 KYDFALYKQDEAFDNKNI PALYSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 301  
 Db 267 KYDFALYKQDEAFDNKNI PALVDALRYGTVVSRDPFKAVTFVANHTDIIWNKYPAY 326  
 QY 302 AFILTYEQOPTIFVRDYEEWLNKDLKNIWIHENLAGGSTDIVYDNDDELIFVRNGYGD 361  
 Db 327 AFILTYEQOPTIFVRDYEEWLNKDLKNIWIHENLAGGSTDIVYDNDDELIFVRNGYGD 386  
 QY 362 KPGIITVINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPA 421  
 Db 387 KPGIITVINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPA 446  
 QY 422 NGQYGSVWSYCGVG 436  
 Db 447 NGYGSVWSYCGVG 461

## RESULT 3

Q8NKR4 Q8NKR4 PRELIMINARY; PRT; 461 AA.  
 ID Q8NKR4  
 AC Q8NKR4  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Alpha-amylase precursor.  
 GN BD5031.  
 OS Thermococcus sp. 'AEP11 1a'.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Thermococcus.  
 OX NCBI\_TaxID=195101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AEP11 1a;  
 RX MEDLINE=22113025; PubMed=11994309;  
 RA Richardson T.H., Tan X., Frey G., Callen W., Cabell M., Lam D.,  
 RA Macomber J., Short J.M., Robertson D.E., Miller C.;  
 RT "A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
 RT AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE.";  
 RL J. Biol. Chem. 277:26501-26507(2002).  
 DR EMBL; AF504063; AA48113.1; -;  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR005975; P:carbohydrate metabolism; IEA.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Amy; 1.  
 KW Signal.  
 FT SIGNAL  
 SQ SEQUENCE 1 26 POTENTIAL.  
 Query Match 92.7%; Score 2279; DB 1; Length 461;  
 Best Local Similarity 91.5%; Pred. No. 4.8e-146;  
 Matches 398; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

QY 2 AKYSELEKGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGWGAYSM 61  
 Db 27 AKYSELEKGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGWGAYSM 86  
 QY 62 GYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWNP 121



```

Db      87  GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGIKVIADIVINHRRAGDLEWN 146
QY      122  PFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWASQ 181
Db      147  PFVGDYTWTFDFSKVASGKYTANYLDFHPNEVNEKCCDEGTFGGPDIAHEKEMDQHWLWASD 206
QY      182  ESYAAALRSIGIDAWRFDVVKGYPVWVKWLNWGGWAGVGYWDTNVDALINWYSSGA 241
Db      207  ESYAAALRSIGIDAWRFDVVKGYPVWVKWLNWGGWAGVGYWDTNVDALINWYSSGA 266
QY      242  KVDFDALYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAATFVANHTDIIWNKYPAY 301
Db      267  KVDFPFLYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAATFVANHTDIIWNKYLAY 326
QY      302  APILTYEGOPTIFRYDYEELWLNKDLNLIWHENLAGGSTDIYVYDNDELIFVRNGYGD 361
Db      327  APILTYEGOPVIFRYDYEELWLNKDLNLIWHENLAGGSTSIYVYDSDEMIYVRNGYGS 386
QY      362  KPGLITYINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 421
Db      387  KPGLITYINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 446
QY      422  NGQYGYVSWSYCGVG 436
Db      447  NGQYGYVSWSYCGVG 461

RESULT 4
C33476 PRELIMINARY; PRT; 461 AA.
ID C33476
AC C33476
NC 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DE Alpha-amylose precursor.
GN APKA.
OS Pyrococcus kodakaraensis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KOD1;
RA Tachibana Y., Mendez L., Fujiwara S., Takagi M., Imanaka T.;
RT "Cloning and expression of the alpha-amylose gene from the
RT hyperthermophilic archaeon Pyrococcus sp. KOD1, and characterization
RT of the enzyme."
RL J. Ferment. Bioeng. 82:224-232(1996).
DR EMBL; D83793; BAR21130.1; -.
DR HSSP; P06278; IVUS.
DR GO; GO:0004556; F:alpha-amylose activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_aml_cat.
DR Pfam; PF00128; Glyco_hydro_13.
DR PRINTS; PR00110; ALPHAAMYLASE.
KW Signal.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 461 AA; 52213 MW; FCC131A93DC03123 CRC64;

```

```

Query Match 91.4%; Score 2248; DB 1; Length 461;
Best Local Similarity 90.6%; Pred. No. 5.9e-144;
Matches 394; Conservative 18; Mismatches 23; Indels 0; Gaps 0;

QY      2  AKYSELEKGGVIMQAFYWDVPSGGIWNDTIRQKIPWYDAGISAIWPPASKMGGGAYSM 61
Db      27  AKYSELEKGGVIMQAFYWDVPSGGIWNDTIRQKIPWYDAGISAIWPPASKMGGGAYSM 86
QY      62  GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGIKVIADIVINHRRAGDLEWN 121
Db      87  GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGIKVIADIVINHRRAGDLEWN 146

```

```

QY      122  PFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWASQ 181
Db      147  PFVGDYTWTFDFSKVASGKYTANYLDFHPNEVNEKCCDEGTFGGPDIAHEKEMDQHWLWASD 206
QY      182  ESYAAALRSIGIDAWRFDVVKGYPVWVKWLNWGGWAGVGYWDTNVDALINWYSSGA 241
Db      207  ESYAAALRSIGIDAWRFDVVKGYPVWVKWLNWGGWAGVGYWDTNVDALINWYSSGA 266
QY      242  KVDFDALYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAATFVANHTDIIWNKYPAY 301
Db      267  KVDFPFLYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAATFVANHTDIIWNKYPAY 326
QY      302  APILTYEGOPTIFRYDYEELWLNKDLNLIWHENLAGGSTDIYVYDNDELIFVRNGYGD 361
Db      327  APILTYEGOPVIFRYDYEELWLNKDLNLIWHENLAGGSTSIYVYDSDEMIYVRNGYGS 386
QY      362  KPGLITYINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 421
Db      387  KPGLITYINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 446
QY      422  NGQYGYVSWSYCGVG 436
Db      447  NGQYGYVSWSYCGVG 461

RESULT 5
O93647 PRELIMINARY; PRT; 457 AA.
ID O93647
AC O93647
NC 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DE Alpha-amylose (EC 3.2.1.1).
GN AMY.
OS Thermococcus hydrothermalis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=46539;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AL662;
RA Leveque E., Haye B., Belarbi A.;
RT "Cloning and expression of an alpha-amylose encoding gene from the
RT hyperthermophilic archaeobacterium Thermococcus hydrothermalis and
RT biochemical characterization of the recombinant enzyme."
RL FEMS Microbiol. Lett. 186:67-71(2000).
DR EMBL; AF068255; AAC97877.1; -.
DR HSSP; P06278; IVUS.
DR GO; GO:0004556; F:alpha-amylose activity; IEA.
DR GO; GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_aml_cat.
DR Pfam; PF00128; alpha-amylose; 1.
DR GlycoSite; Hydrolyase.
SQ SEQUENCE 457 AA; 51543 MW; 602367643AC7847B CRC64;

```

```

Query Match 90.3%; Score 2221; DB 1; Length 457;
Best Local Similarity 89.9%; Pred. No. 3.9e-142;
Matches 391; Conservative 20; Mismatches 24; Indels 0; Gaps 0;

```

```

QY      2  AKYSELEKGGVIMQAFYWDVPSGGIWNDTIRQKIPWYDAGISAIWPPASKMGGGAYSM 61
Db      23  AKATLENGGVIMQAFYWDVPSGGIWNDTIRQKIPWYDAGISAIWPPASKMGGGAYSM 82
QY      62  GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGIKVIADIVINHRRAGDLEWN 121
Db      83  GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGIKVIADIVINHRRAGDLEWN 142
QY      122  PFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWASQ 181
Db      143  PFTNSYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQHWLWASN 202

```

QY 182 ESVAAYLRISIGIDAMRFDYKGYAPVWVVDLNNWGGWAVGEBYDNTVDVNLWAYSSGA 241  
 Db 203 ESVAAYLRISIGIDAMRFDYKGYAPVWVVDLNNWGGWAVGEBYDNTVDVNLWAYSSGA 262  
 QY 242 KVFDFALYKMDAFDNKNI PALVSALONGQTVVSRDPKAVTFVANHTDIIWNKYPAY 301  
 Db 263 KVFDFALYKMDAFDNKNI PALVSALONGQTVVSRDPKAVTFVANHTDIIWNKYPAY 322  
 QY 302 AFILTYEGOPTFYRDEEWNKDLKNIWIHNLNLAGGSTDIVYDNDDELIFVRNGYD 361  
 Db 323 AFILTYEGOPTFYRDEEWNKDLKNIWIHNLNLAGGSTDIVYDNDDELIFVRNGYD 382  
 QY 362 KPLIITYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 421  
 Db 383 KPLIITYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 442  
 QY 422 NGQGYVMSYCGVG 436  
 Db 443 NGQGYVMSYCGVG 457  
 RESULT 6  
 O50200 PRELIMINARY; PRT; 469 AA.  
 ID O50200  
 AC O50200  
 DT 01-JUN-1998 (T-EMBLrel. 06, Created)  
 DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
 DE Amylase (EC 3.2.1.1).  
 GN AMY.  
 OS Thermococcus sp. Rt3.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Thermococcus.  
 OX NCBI\_TaxID=65421;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rt3;  
 RX MEDLINE=99154282; PubMed=10030014;  
 RA Jones R.A., Jermin L.S., Eastal S., Patel B.K., Beacham I.R.;  
 RT "Amylase and 16S rRNA genes from a hyperthermophilic  
 archaeobacterium";  
 RL J. Appl. Microbiol. 86:93-107(1999).  
 DR EMBL; AF017454; AAB87860.1; -;  
 DR HSP; P06278; 1VJS.  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha-amyl cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 469 AA; 52756 MW; C3BD614FBA4DB3E2 CRC64;  
 Query Match 89.4%; Score 2199.5; DB 1; Length 469;  
 Best Local Similarity 50.5%; Pred. No. 1.1e-140;  
 Matches 389; Conservative 16; Mismatches 24; Indels 1; Gaps 1;  
 QY 7 LEKGGVIMQAFYVDVSGGIWDTIRQKIPWYDAGISAIWIPPSKMGAYSGYDPPY 66  
 Db 41 LEKGGVIMQAFYVDVSGGIWDTIRQKIPWYDAGISAIWIPPSKMGAYSGYDPPY 100  
 QY 67 DFDLGEYDQKGTETRFSGKQLVNMINTAHAYGMKVADIVINHRAGDLEWPFVND 126  
 Db 101 DFDLGEYDQKGTETRFSGKQLVNMINTAHAYGMKVADIVINHRAGDLEWPFVND 160  
 QY 127 YTWDFSKVAGKYTANYLDHFNELHAGSGTFGGYDPICHDKSWDQYWLWASQESVAA 186  
 Db 161 YTWDFSKVAGKYTANYLDHFNELHAGSGTFGGYDPICHDKSWDQYWLWASQESVAA 220  
 QY 187 YLRSIGIDAMRFDYKGYAPVWVVDLNNWGGWAVGEBYDNTVDVNLWAYSSGA 246  
 Db 221 YLRSIGIDAMRFDYKGYAPVWVVDLNNWGGWAVGEBYDNTVDVNLWAYSSGA 279

QY 247 ALYKMDAFDNKNI PALVSALONGQTVVSRDPKAVTFVANHTDIIWNKYPAYAFILT 306  
 Db 280 ALYKMDAFDNKNI PALVSALONGQTVVSRDPKAVTFVANHTDIIWNKYPAYAFILT 339  
 QY 307 YEGOPTFYRDEEWNKDLKNIWIHNLNLAGGSTDIVYDNDDELIFVRNGYD 366  
 Db 340 YEGOPTFYRDEEWNKDLKNIWIHNLNLAGGSTDIVYDNDDELIFVRNGYD 399  
 QY 367 TYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 426  
 Db 400 TYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 459  
 QY 427 YSVMSYCGVG 436  
 Db 460 YSVMSYCGVG 469  
 RESULT 7  
 O08452 PRELIMINARY; PRT; 460 AA.  
 ID O08452  
 AC O08452  
 DT 01-JUL-1997 (T-EMBLrel. 04, Created)  
 DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Alpha amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase)  
 GN AMYA.  
 OS Pyrococcus furiosus, and  
 OS Pyrococcus woesei.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=2261, 2262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.furiosus; STRAIN=DSM 3638;  
 RA Dong G., Vieille C., Savchenko A., Zelkus J.G.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.furiosus; STRAIN=DSM3638;  
 RX MEDLINE=97341170; PubMed=9195939;  
 RA Jorgensen S., Vorgias C.E., Antranikian G.;  
 RT "Cloning, sequencing, characterization, and expression of an  
 extracellular alpha-amylase from the hyperthermophilic archaeon  
 Pyrococcus furiosus in Escherichia coli and Bacillus subtilis";  
 RL J. Biol. Chem. 272:16335-16342(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.furiosus; STRAIN=DSM3638;  
 RA Jorgensen S.T., Vorgias C.E., Antranikian G.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.woesei; STRAIN=DSM3773;  
 RA Chunlin Lu, Weizheng J., Yunyan Y.;  
 RT "Cloning and Expression of Alpha Amylase from Pyrococcus woesei";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 DR EMBL; AF001268; AAC45663.1; -;  
 DR EMBL; U96622; AAB67705.1; -;  
 DR EMBL; AF177906; AAD54338.1; -;  
 DR HSP; P06278; 1VJS.  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha amyl cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 460 AA; 52909 MW; 0E0A2AEB5FCAF541 CRC64;

Query Match 87.6%; Score 2153; DB 1; Length 460;  
Best Local Similarity 86.7%; Pred. No. 1.6e-137;  
Matches 377; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

2 AKYSELEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGSGGYSM 61  
26 AKYLEEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGSGGYSM 85  
62 GYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWN 121  
86 GYDYPDFDLGEYDQKGTETRFSGKQELVRLIQTAHAYGIKVIADIVINHRAGGDLWN 145  
122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTFFGGYDICHDKSWDQYWLWASQ 181  
146 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHCCDEGTFFGGYDICHKKEWDQYWLWKS 205  
182 ESYAAYLRSIGDWARFDYKGYAPWVVKDMLNMGWNAVGEYDNTDNLWVAYSSGA 241  
206 ESYAAYLRSIGDWARFDYKGYAPWVVKDMLNMGWNAVGEYDNTDNLWVAYSSGA 265  
242 KVFDFALYKMDAEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 301  
266 KVFDFALYKMDAEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 325  
302 AFILTYEGOPTIFRYDYEELNKKLNLINWHENLAGSGTDIVYDNDDELIFVRNGYGD 361  
326 AFILTYEGOPTIFRYDYEELNKKLNLINWHENLAGSGTDIVYDNDDELIFVRNGYGD 385  
362 KPGLITYINLSSKAGRWVYVVKFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPYDPA 421  
386 KPGLITYINLSSKAGRWVYVVKFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPYDPA 445  
422 NGQYGSVMSYCGVG 436  
446 NGYIGYSVMSYCGVG 460

RESULT 8  
Q8U319 PRELIMINARY; PRT; 473 AA.  
AC Q8U319;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-amylase.  
GN PF0477.  
OS Pyrococcus furiosus.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OX Pyrococcus.  
OX NCBI\_TaxID=2261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;  
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
RT "The complete sequence of the Pyrococcus furiosus genome.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB010170; AAL80601.1;  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amy\_cat.  
DR InterPro; IPR006046; Glyco\_hydro\_13.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00110; ALPHAAMYLASE.  
KW Complete proteome.  
SQ SEQUENCE 473 AA; 54394 MW; 2E212A2C9A67F3B5 CRC64;

Db 39 AKYLEEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGSGGYSM 98  
QY 62 GYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWN 121  
Db 99 GYDYPDFDLGEYDQKGTETRFSGKQELVRLIQTAHAYGIKVIADIVINHRAGGDLWN 158  
QY 122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTFFGGYDICHDKSWDQYWLWASQ 181  
Db 159 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHCCDEGTFFGGYDICHKKEWDQYWLWKS 218  
QY 182 ESYAAYLRSIGDWARFDYKGYAPWVVKDMLNMGWNAVGEYDNTDNLWVAYSSGA 241  
Db 219 ESYAAYLRSIGDWARFDYKGYAPWVVKDMLNMGWNAVGEYDNTDNLWVAYSSGA 278  
QY 242 KVFDFALYKMDAEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 301  
Db 279 KVFDFALYKMDAEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 338  
QY 302 AFILTYEGOPTIFRYDYEELNKKLNLINWHENLAGSGTDIVYDNDDELIFVRNGYGD 361  
Db 339 AFILTYEGOPTIFRYDYEELNKKLNLINWHENLAGSGTDIVYDNDDELIFVRNGYGD 398  
QY 362 KPGLITYINLSSKAGRWVYVVKFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPYDPA 421  
Db 399 KPGLITYINLSSKAGRWVYVVKFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPYDPA 458  
QY 422 NGQYGSVMSYCGVG 436  
Db 459 NGYIGYSVMSYCGVG 473

RESULT 9  
Q9P9L0 PRELIMINARY; PRT; 460 AA.  
AC Q9P9L0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-amylase.  
OS Pyrococcus woesei.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OX Pyrococcus.  
OX NCBI\_TaxID=2262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang C., Lo H.;  
RT "Cloning of Pow alpha-amylase gene.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF240464; AAF44693.1;  
DR HSSP; P06278; 1VJS.  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amy\_cat.  
DR InterPro; IPR006046; Glyco\_hydro\_13.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00110; ALPHAAMYLASE.  
SQ SEQUENCE 460 AA; 53008 MW; C6734D3B8578A966 CRC64;

Query Match 87.5%; Score 2152; DB 1; Length 460;  
Best Local Similarity 86.7%; Pred. No. 1.8e-137;  
Matches 377; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

2 AKYSELEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGSGGYSM 61  
26 AKYLEEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGSGGYSM 85  
62 GYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGGDLWN 121  
86 GYDYPDFDLGEYDQKGTETRFSGKQELVRLIQTAHAYGIKVIADIVINHRAGGDLWN 145  
122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTFFGGYDICHDKSWDQYWLWASQ 181  
146 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHCCDEGTFFGGYDICHKKEWDQYWLWKS 205

QY 182 ESYAAYLSIGIDAWRFYKGYAPVYVYKDLNWMGAGVGEYWDNTVDALNWAYSSGA 241  
 DB 206 ESYAAYLSIGIDGWFYKGYAPVYVYKDLNWMGAGVGEYWDNTVDALNWAYSSGA 265  
 QY 242 KYDFDALYKMDKDEAFDKNIPALYSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPAY 301  
 DB 266 KYDFDALYKMDKDEAFDKNIPALYSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPAY 325  
 QY 302 AFILTYEQOPTIFRYDYEEWLNKDLNLIWIHENLAGGSTDIWYDNDDELIFVRNGD 361  
 DB 326 AFILTYEQOPTIFRYDYEEWLNKDLNLIWIHENLAGGSTDIWYDNDDELIFVRNGD 385  
 QY 362 KGLIYINLSSKAGRWVYVYKAGACIHEYTCNLCGWVYKGYSGWVYLEAPAYDPA 421  
 DB 386 KGLIYINLSSKAGRWVYVYKAGACIHEYTCNLCGWVYKGYSGWVYLEAPAYDPA 445  
 QY 422 NGYGYVSWYSCGVG 436  
 DB 446 NGYGYVSWYSCGVG 460

## RESULT 10

Q987H6 PRELIMINARY; PRT; 519 AA.  
 ID Q987H6  
 AC Q987H6  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Raw starch digesting amylase precursor.  
 OS Cytophaga sp.  
 OC Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales;  
 OC Flexibacteraceae; Cytophaga.  
 OX NCBI\_TaxID=29535;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jeang C.L., Chen L.S., Chen M.Y.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF067653; AAF00567.1; -  
 DR HSSP; P06278; 1VUS.  
 DR GO; GO:0004556; P:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha amyl cat.  
 DR InterPro; IPR006589; Alp amyl cat sub.  
 DR InterPro; IPR006046; Glyco hydro 13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Amy; 1.  
 KW Signal.  
 FT SIGNAL.  
 FT CHAIN  
 SQ SEQUENCE 519 AA; 58337 MW; 3E6B88A4DF98B163 CRC64;

Query Match 20.9%; Score 513; DB 2; Length 519;  
 Best Local Similarity 31.1%; Pred. No. 1.2e-26;  
 Matches 160; Conservative 64; Mismatches 159; Indels 132; Gaps 24;  
 QY 11 GVIMQAFYWDVPSGGIWWDTIRQIPWYDAGISAIWIPPASKMGCGAYSMGYDPPDF 70  
 DB 39 GTWQYFEWVPNDGQWNRRLTDAPYLSVGITAVTPPAYKTSQA-DVGYGPDLYD 97  
 QY 71 LGEYDQGTETVTRFGSKQELVNNINTAHAYGMKVIADIVINHRAGD-----LEWNP 123  
 DB 98 LGEFNQGTVRTKGTGELKLSAINTLHNSGIGVYGDVYVFNHRGGADGTENYAVEVNS 157  
 QY 124 -----VNDYTWDTDFSKVASGKYTANYLDLPHNELH----- 153  
 DB 158 NRNQETSGEYNIQAWTGFNPPGRG-----TTSNFKWQWPFHFDGTQDQDQSLRIFKRG 214  
 QY 154 -----AGDSGTFG-----GYPDICHD-KSWDQYWLWASQESYAYLRIGIDA 195  
 DB 215 GKAWDWEVSSGENGYDLYMADIDYDHPDVPVNMKKGW---VW-----YANEVGLDG 263

QY 196 WRFDYKGYAPVYVYKDLNWMGAGVGEYWDNTVDALNWAYSSGA 243  
 DB 264 YRLDAVGHKIFSLKQWVDNARAATKEMFTVGEYQWQDGLNLYL-----AKVYVNSQL 319  
 QY 244 FDPALYKMDKDEAFDKNIPALYSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPAY 292  
 DB 320 FDPALYKMDKDEAFDKNIPALYSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPAY 371  
 QY 293 -----IWNKYPAYAFILTYE-GQTFITFYRDY-----EELNKKDLNLIWIHENLA 338  
 DB 372 ESTVQPMFKELAYAFILTYE-GQTFITFYRDY-----EELNKKDLNLIWIHENLA 431  
 QY 339 GGSFDIYVYDNDLFI-FVRNGYCD--KPLIYINLSSKAGRWVYVYKAGACIHEYTCN 395  
 DB 432 YG-TQDYIINPDIWGTREGDSTKAKSLATVITDPPGSGKMYVGTSGNAGEIWIYDUTG 490  
 QY 396 NLGSGWVYKGYSGWVYLEAPAYDPPANGQYGSVW 430  
 DB 491 NR---TDKITIGS-----DGYATFPVNGG-SVSVW 516  
 RESULT 11  
 Q987H6 PRELIMINARY; PRT; 507 AA.  
 ID Q987H6  
 AC Q987H6  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Alpha-amylase.  
 GN VPA0999.  
 OS Vibrio parahaemolyticus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=RMD 2210633 / Serotype O3:K6;  
 RX MEDLINE=22508454; PubMed=12620739;  
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
 distinct from that of V. cholerae";  
 RL Lancet 361:743-749(2003).  
 DR EMBL; AP005087; BAC62342.1; -  
 DR GO; GO:0004556; P:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha amyl cat.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 507 AA; 58428 MW; 91B549E2BE0700BD CRC64;

Query Match 20.8%; Score 511.5; DB 16; Length 507;  
 Best Local Similarity 27.5%; Pred. No. 1.5e-26;  
 Matches 143; Conservative 81; Mismatches 157; Indels 139; Gaps 23;  
 QY 9 KGVIMQAFYWDVPSGGIWWDTIRQIPWYDAGISAIWIPPASKMGCGAYSMGYDPPDF 68  
 DB 25 QNGTWQYFEWVPNDGALWQVESNAPALAEENGFTALWLPAYKAGSGNDVGYVDM 84  
 QY 69 FDLGEYDQGTETVTRFGSKQELVNNINTAHAYGMKVIADIVINHRAGD-----LE 119  
 DB 85 YDLGEFDQKSVRTKGTGKQAQYISAINAAHNNNTIYIGDVVFNHRGGADGKSWDYTRKD 144  
 QY 120 WN-----PFVND-----YTWTFDS----- 133  
 DB 145 WDNRNIELGDKWIEAWVEFPGRNDKYSNFHWTFYHFDGVDWDDAGKEAIKFKGEGK 204  
 QY 134 -----KVASKYTYANVLDLPHNELHAGDGTGGYDPCIDHD-KSWDQYWLWASQESYAYL 188  
 DB 205 ANDWEVSSKGYNDYLYMADLDM-----DHFVQKELKDWGEW-----YI 244  
 QY 189 RSIGIDAWRFYKGYAPVYVYKDLNWMGAGVGEYWDNTVDALNWAYSSGA 240

Db 245 NMTGVDFRMDAVKHLIKYQLQEWIDHLRWKTKGELFTVGEYWNVDVNLHNFINTSGS 304  
QY 241 AKVDFEAL---YKXDEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTII--- 294  
Db 305 MSDFDAPLHNFNFNASKSGNYD---MRQIMNG-TLMKDNFVKAVTLVNHNDTQLOAL 359  
QY 295 -----WNKYPAYAFI-LTYEGQPTIFRYDYEWLKDK-----LKNLIWIHEN 336  
Db 360 ESTVDWFKPLAYAFILLREEGVPSFYADYGAQYSDKGYNNMAKVPYIBELVTLRKE 419  
QY 337 LAGGSTDIVYNDDEL-FVRNGYDKPGLIYINLGSSKAGRWYVVPKAGACIHEYTG 395  
Db 420 YAYGQN-SYLDHWDVIGWTRGDADHFNMAVIMSDGPGTKMY-----TGKPSRYVD 474  
QY 396 NLG-----GWVDKYVYSSGWVYLEAPAYDPANGQYGVSW 430  
Db 475 KLGIRTEVWTD-----ANGW-----AEPVNGG-SVSW 503  
RESULT 12  
Q8YU21 PRELIMINARY; PRT; 492 AA.  
AC Q8YU21  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Alpha-amylase.  
GN ALR2190.  
OS Anabaena sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium Anabaena sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
DR EMBL; AP003588; BAB73889.1; -  
DR PIR; AH2079; AH2079  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR InterPro; IPR006046; Glyco\_hydro\_13.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00110; ALPHAAMYLASE.  
DR Complete Proteome.  
KW SEQUENCE 492 AA; 57063 MW; 11B5D6ECP2F18288 CRC64;  
SQ SEQUENCE 492 AA; 57063 MW; 11B5D6ECP2F18288 CRC64;  
Query Match 20.0%; Score 492; DB 16; Length 492;  
Best Local Similarity 27.8%; Pred. No. 3.1e-25;  
Matches 147; Conservative 81; Mismatches 153; Indels 148; Gaps 28;  
QY 8 EKGGVIMQAFVWDVPSGGIWDTRQKIPEDYDAGISAIMPPASKMGAGYNGYDPYD 67  
Db 3 QMGNTMQLFYHWYFNDGNLMSKVEASAPELADAGFTAMWLPFAKGFAGSFDVGVGYD 62  
QY 68 PFDLGEYDQKGTVTRFGSKOELVNMINTAHAYGMKVIADIVNHAGGDL-----EWNPF 123  
Db 63 LFDLGEFDQKSVTKYTRQYLDVAKSLQTHGLQYVADAVLNHRKGGGAVETPKATPF 122  
QY 124 VND-----YTWDTSKVASGKYTANY-----LDFHPNHLHAGDSGT- 159  
Db 123 PQDRLNPKGGLQDIKITYTHNPPG-RQGYK-SNFEWHWHFADVDY-NEYNSGDRSTV 178  
QY 160 -----EGGYP-----DICHKSDQ-----YWLWASOESVAAYLRSIGID 194  
Db 179 YLLSGKNFDDYVALEKGNFAYLMGCCDLDFQNEWRVGEVTVWGRWC-----LDTTKVD 230

QY 195 AWRTDYKGVAPVWVKNLWNGGWA-----VGEYWDITNDVAVLNW-----AYSSGAKVFD 245  
Db 231 GFRIDAIKIHISTWTFPEWIDALESHAGKDLFMVGEYWNIDINTLL-WVDVAVRKMVSVD 289  
QY 246 FALYKXDEA-----FDKNIPALVSALONGQTVVSRDPFKAVTFVANHDTII----- 294  
Db 290 VFLHYNFHQASKSGNVDMMRI-----LDG-TMMQORPHTAVTFVENHDSQPLQALES 341  
QY 295 -----WNKYPAYAFI-LTYEGQPTIFRYD-----YEWLKDK-----LKNLI 331  
Db 342 VVEWFKPLAYAFILLREEGVPCVFHADYGAEDW-GKDGNNYINFPSHRWIDKLL 400  
QY 332 WIHENLAGGSTDIVYNDDEL-FVRNGYDKPGLIYINLGSSKAGRWYV-----PKPA 386  
Db 401 YARKHAYG-PQNYLDHWNITGWTRGLGDADHPQGMVIMSDGSEGIKMWVEGKPKTKFI 459  
QY 387 GACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPANGQY-----GYSVW 430  
Db 460 DLTEH-----IKEAVTYNEWG-----GEFRLGGSVSVW 489  
RESULT 13  
Q8Z5S5 PRELIMINARY; PRT; 494 AA.  
AC Q8Z5S5  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Cytoplasmic alpha-amylase.  
GN STY2171 OR AMYA OR T0914.  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=601;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21534947; PubMed=11677608;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
RA Feltwell T., Hamlin N., Haque R.M., Dowd L., White N., Farrar J.,  
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,  
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
RA Whitehead S., Barrrell B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ty2 / ATCC 700931;  
RX MEDLINE=22531367; PubMed=12644504;  
RA Dang W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
RA Burland V., Kodyouranni V., Schwartz D.C., Blattner F.R.;  
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
and CT18.";  
RL J. Bacteriol. 185:2330-2337(2003).  
RN EMBL; AL627272; CAD05711.1; -  
DR EMBL; AS016837; AAO68592.1; -  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR InterPro; IPR006589; Alp amyl cat\_sub.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Amy; 1.  
DR Complete proteome.  
KW SEQUENCE 494 AA; 56451 MW; 5D221A7B24F67D4 CRC64;  
SQ SEQUENCE 494 AA; 56451 MW; 5D221A7B24F67D4 CRC64;  
Query Match 19.8%; Score 487.5; DB 16; Length 494;  
Best Local Similarity 28.7%; Pred. No. 6.2e-25;  
Matches 139; Conservative 71; Mismatches 156; Indels 119; Gaps 21;

QY 9 KGGVIMQAFYMDVPSGGIWMDTIRQKIPWYDAGISAIWIPBASKMGAGYSGYDYPDF 68  
 DB 2 KNPITLQVHFHYFDPGGKLNSELAEADGLNDIGINWVLPBCKGASGYSYGYDSYDL 61  
 QY 69 FDLGSDQKGTETRFSGKQBLVNNINTAHAYGMKVIADIVINHRAGDLENNPFV 124  
 DB 62 FDLGSDQKGTETRFSGKQBLVNNINTAHAYGMKVIADIVINHRAGDLENNPFV 121  
 QY 125 NDYR-----WTFDSKVA-SGKYTANYLDHF-----BNELHAGDSGTF- 160  
 DB 122 QDRTQIDNLIIBECWTRITFFPARAGYSNFIWDYHCFSGIDHIENPDE-----DGIFK 176  
 QY 161 -----GGYDICHK--SWD-----QYWL-WASQESYAAYLRSIG 192  
 DB 177 IVNDYTGDMNDQVDELGNFDYLMGENIDFRNHAVTEEEKYWARVWMEQTHC----- 229  
 QY 193 IDAMFDFYVKGYPAPWVKDLNMGWGA-----VGEYDNTVDVNLWAYSSGAK--VF 244  
 DB 230 -DGFRLDAVKHIDPAWFKYKWEHIEHVQAVAPKPLFIVAIEYSHVEDVKQTYIDQDQKTMUF 288  
 QY 245 DFALYKMDA-----FDKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTII----- 294  
 DB 289 DAPLQMKFHEASROGAEYDMEHIFT-----GTLVEADPFPAVTLVANHDTQLLQALE 340  
 QY 295 -----WNKYPAVAFILTYE-GOPTIFERYD-----EVLNKD-----KLNLIW 332  
 DB 341 APVEWFKPLAYALLIRENGVPSVFPDLYGASIEDNGENGTCRVDMPVINQLDRLIL 400  
 QY 333 IHENLAGSDTIVYDNDLEI-FVRNGYDKPGLITVINLGSSKAGRWVVPKFAAGACIH 391  
 DB 401 ARQRFPAHG-IQTLFDPHPCIAFSRSGTEENPGCVVLSNGDDGCKTLILGDNVANKTW 459  
 QY 392 EYTG 396  
 DB 460 DFLGN 464

RESULT 14  
 Q93148  
 ID Q93148 PRELIMINARY; PRT; 501 AA.  
 AC Q93148  
 DT 01-DEC-2001 (TremBLrel. 19, Created)  
 DT 01-DEC-2001 (TremBLrel. 19, last sequence update)  
 DT 01-JUN-2003 (TremBLrel. 24, last annotation update)  
 DE Amylase.  
 GN AMYK38.  
 OS Bacillus sp. KSM-K38.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=129736;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KSM-K38;  
 RA Hayashi Y.;  
 RT "Isolation of a new Bacillus alpha-amylase.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB051102; BAB71820.1; -  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha\_amyl\_cat.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR SMART; SM00642; Aamy; 1.  
 SQ SEQUENCE 501 AA; 57485 MW; 1240F46739A5CC11 CRC64;

Query Match 19.8%; Score 486; DB 2; Length 501;  
 Best Local Similarity 30.28; Pred. No. 8e-25;  
 Matches 149; Conservative 55; Mismatches 163; Indels 126; Gaps 24;  
 QY 11 GVIMQAFYMDVPSGGIWMDTIRQKIPWYDAGISAIWIPBASKMGAGYSGYDYPDF 70  
 DB 26 GTMYYEWHLENDQGHNLHDDAAALSAGITAIWIPPAYKNSQA-DVGYGAYDLYD 84  
 QY 71 LGEYDQKGTETRFSGKQBLVNNINTAHAYGMKVIADIVINHRAGDLENNPFV 118

DB 85 LGEFNQKGTVRTYKTKAQLERAIKSLKSNIDINVGVVNNHKGADFTTEAVQAVQVNP 144  
 QY 119 -EWNPFNDYT---WTFDSKVASGKYTANYLDHFPHNELHAGSGTGGYDPICHDKSWDQ 174  
 DB 145 NRWODISGAYTIDAWTGFD--FSGRNA--YSDFKRWFEH-----FNGV-----DWDQ 188  
 QY 175 -----YWLWASQESYAAYLRSIG-----IDA 195  
 DB 189 RYQENHIFRPANTNNWVDEENGNYDYLGNISIDFSPHEVQDELKMGSMFTDELDDG 248  
 QY 196 WRFDYVKGYPAPWVKDLNMGWGA-----WAVEYDNTVDV-----LNWAYSSGAKV 243  
 DB 249 YRLDAIKHIFWYTSMDVRHNEADQDLFVGEYKWDVGALEFYLDENMWEMS---L 304  
 QY 244 DFALYKMDA-----FDKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTI----- 292  
 DB 305 FDVPLNIFYEASQOQGSYDMRNI-----LRG--SLVEAHPMHAFTFVDNHDTPGESL 356  
 QY 293 -----IWNKYPAVAFILTYE-GOPTIFERYDYEELN-----KDKLNLINHELAGS 341  
 DB 357 ESWVADWFKPLAYATILTRGGYPNVFYGYGIPNDNISAKKMDIDELLDARQYAYG- 415  
 QY 342 TDIVYDNDLEI-FVRNGYDKP--GLITVINLGSSKAGRWVVPK--PAGACIHEVYTNL 397  
 DB 416 TQHDYFDHVDVGTREGSSSRFNSGLATIMSGPG--GSKWYVGRQAGQWTDLTGN 474  
 QY 398 GGWDYKVIYSSGW 410  
 DB 475 GASVT--INGDGV 485

RESULT 15  
 O82839  
 ID O82839 PRELIMINARY; PRT; 516 AA.  
 AC O82839  
 DT 01-NOV-1998 (TremBLrel. 08, Created)  
 DT 01-NOV-1998 (TremBLrel. 08, last sequence update)  
 DT 01-JUN-2003 (TremBLrel. 24, last annotation update)  
 DE Amylase.  
 OS Bacillus sp.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1409;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KSM-1378;  
 RA Igashiki K., Hatada Y., Ikawa K., Araki H., Ozawa T., Kobayashi T.,  
 RA Ozaki K., Ito S.;  
 RT "Improved thermostability of a Bacillus alpha-amylase by deletion of  
 RT an arginine-glycine residue is caused by enhanced calcium binding.";  
 RL Biochem. Biophys. Res. Commun. 248:372-377(1998).  
 DR EMBL; AB008763; BAA32431.1; -  
 DR HSP; P06278; 1VUS.  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha\_amyl\_cat.  
 DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Aamy; 1.  
 SQ SEQUENCE 516 AA; 58841 MW; D90A8C90ECC182F8 CRC64;

Query Match 19.2%; Score 472.5; DB 2; Length 516;  
 Best Local Similarity 29.9%; Pred. No. 6.8e-24;  
 Matches 147; Conservative 66; Mismatches 156; Indels 119; Gaps 25;  
 QY 11 GVIMQAFYMDVPSGGIWMDTIRQKIPWYDAGISAIWIPBASKMGAGYSGYDYPDF 70  
 DB 38 GTMYYEWHLENDQGHNLHDDAAALSAGITAIWIPPAYKNSQA-DVGYGAYDLYD 96  
 QY 71 LGEYDQKGTETRFSGKQBLVNNINTAHAYGMKVIADIVINHRAGDLENNPFV 121

Db 97 LGFPMKGTVRKYGTRSLQCAVTSLKXNGIGQVIGDVVMNKHKGADGTEMWNAVEVNRS 156  
QY 122 ----PFVNDYT---WTFDSKVASGYTANY-----LDFPHELHAGDSGTFGG--- 162  
Db 157 NRNQBISGEYTIETAWTKEDFPGRGNTHSNFKRWYHFDGTDWDQSRQLQNKIYKFRGTGK 216  
QY 163 -----YPDICHD-----KSWDQYWLNASQBSYAAYLRSIGIDAWR 197  
Db 217 ANDWEVDIENGNYDILMYADIIMDHPEVINELRNG---VW-----YNTLNLDOGR 265  
QY 198 FDIYKGYAPWVVKDWL-----NWGG--WAVEGYWDTNVDVNL-----WAYSSGAKVFD 245  
Db 266 IDAVKHIKYSYTRDMLTHVRNTTGKPMFAVAFWKNDLAAIENYLKNTSWNHS-----VFD 321  
QY 246 FALYYKMDXA-----FDKNIPALVSALONGQTVVSRDPPKAVTFVANHTD----- 292  
Db 322 VPLHYNLYNASNSGGYFDMRNI-----LNG-SYVQKHPITHAVTFVDNHDSPGEALES 373  
QY 293 --IIWNKYPAYAFILTYE-GOPTIFYRDYEEWLN-----KDKLKNLIWIHENLAGGSTD 343  
Db 374 FVQSWFKPLAYALILITREQYPSVYGYGIPTHGVPSMKSIDPILLOARQTYAYG-TQ 432  
QY 344 IVYNDNELI-FVRNGYGDKP--GLITYINLGSSKAGRWVYVPKP-AGACIHEVTGNLGG 399  
Db 433 HDYFDHDIIGWTREGDSSHNSGLATIMSDPG-GNKWMYVGHKAGQVWRDITGNRSG 491  
QY 400 WVDKXVYSSGW 410  
Db 492 TVT--INADGW 500

Search completed: June 29, 2004, 09:35:44  
Job time : 49 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:33:12 ; Search time 22 Seconds  
(without alignments)  
1023.133 Million cell updates/sec

Title: us-10-081-739a-2

Perfect score: 2459

Sequence: 1 NAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVNSYGVG 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	488	19.8	485	2	US-08-446-803-2
2	488	19.8	485	2	US-08-861-837-2
3	488	19.8	485	3	US-08-600-656-2
4	488	19.8	485	3	US-09-170-670-2
5	488	19.8	485	3	US-09-170-670-8
6	488	19.8	485	3	US-09-133-068-2
7	488	19.8	485	3	US-09-193-068-8
8	488	19.8	485	3	US-09-183-412-2
9	488	19.8	485	3	US-09-183-412-8
10	488	19.8	485	3	US-09-264-097-5
11	488	19.8	485	3	US-09-354-191A-2
12	488	19.8	485	4	US-09-291-023A-4
13	488	19.8	485	4	US-09-291-023A-14
14	488	19.8	485	4	US-09-290-734-2
15	488	19.8	485	4	US-09-230-734-8
16	488	19.8	485	4	US-09-381-687-3
17	488	19.8	485	4	US-09-545-586-2
18	488	19.8	485	4	US-09-545-586-8
19	488	19.8	485	4	US-09-540-715A-4
20	488	19.8	485	4	US-09-540-715A-14
21	488	19.8	485	4	US-09-769-864-2
22	488	19.8	485	4	US-09-769-864-8
23	486	19.8	501	4	US-09-465-519-4
24	484	19.7	485	4	US-09-291-023A-2
25	484	19.7	485	4	US-09-291-023A-15
26	484	19.7	485	4	US-09-540-715A-2
27	484	19.7	485	4	US-09-540-715A-15

28 478.5 19.5 483 2 US-08-600-908A-13 Sequence 13, Appl  
29 478.5 19.5 483 3 US-08-683-838A-13 Sequence 13, Appl  
30 478.5 19.5 483 4 US-09-636-252A-13 Sequence 13, Appl  
31 475.5 19.3 515 4 US-09-381-687-7 Sequence 7, Appl  
32 474 19.3 501 4 US-09-465-519-2 Sequence 2, Appl  
33 473.5 19.3 560 3 US-08-814-052-6 Sequence 6, Appl  
34 473.5 19.3 560 3 US-08-812-829-6 Sequence 6, Appl  
35 472.5 19.2 480 3 US-09-182-859-4 Sequence 4, Appl  
36 472.5 19.2 480 3 US-09-170-670-5 Sequence 4, Appl  
37 472.5 19.2 480 3 US-09-193-068-5 Sequence 5, Appl  
38 472.5 19.2 480 3 US-09-183-412-5 Sequence 5, Appl  
39 472.5 19.2 480 4 US-09-290-734-5 Sequence 5, Appl  
40 472.5 19.2 480 4 US-09-672-459-4 Sequence 4, Appl  
41 472.5 19.2 480 4 US-09-545-586-5 Sequence 5, Appl  
42 472.5 19.2 480 4 US-10-186-042-4 Sequence 5, Appl  
43 472.5 19.2 480 4 US-09-769-864-5 Sequence 4, Appl  
44 472.5 19.2 483 4 US-09-291-023A-16 Sequence 16, Appl  
45 472.5 19.2 483 4 US-09-537-168-6 Sequence 6, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-446-803-2  
; Sequence 2, Application US/08446803  
; Patent No. 5824531

; GENERAL INFORMATION:

; APPLICANT: Ctrup, Helle

; APPLICANT: Bisgard-Frantzen, Henrik

; APPLICANT: Ostergaard, Peter Rahbek

; APPLICANT: Rasmussen, Michael Dolberg

; APPLICANT: Van Der Zee, Pia

; TITLE OF INVENTION: Alkaline Bacillus Amylase

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 58245310 No. 5824531disk of No. 5824531th America

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,803

; FILING DATE: 01-June-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Harrington, James J.

; REGISTRATION NUMBER: 38,711

; REFERENCE/DOCKET NUMBER: 4157.204-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 867-0123

; TELEFAX: (212) 878-9655

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 485 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-446-803-2

Query Match 19.8%; Score 488; DB 2; Length 485;

Best Local Similarity 31.0%; Pred. No. 1.1e-34;

Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GYIMQAFYWDVPSGGIMWDTIRQKIPWYDAGISAIWIPPSKMGWGAYSGYDYPDFD 70

Db 7 GTWQYFPEWHLNDGNHNRRLDDASNLNRGITAIWIPAWKGT-SQNDVGYGAYDLYD 65  
QY 71 LGEDYDQKGTETRGSKQELVNMINTAHAYGMKVADIIVINHRAGGD-----LEWNP- 122  
Db 66 LGDFNOKGTETRTKGTGRSQAISALHAKNNGVQVYGVVNMHKGADATENVLAVEVNP 125  
QY 123 -----FVNDYT---WTFDSKVASKYTANYLDHFHNPHELHAGDGTGFGYPDICHDKSWDQ 174  
Db 126 NRNQISGDYITAEATKDFPGRGN---TYSDFKWRWYH-----FDGV-----DWDQ 169  
QY 175 -----YWLWASQES-----YAA-----YLRSIG 192  
Db 170 SRQFNRIYKFRGKAWDEVDSENGNYDYLMDYADVMDHPEVNVNLRWGEWYNTLN 229  
QY 193 IDAWRFDYVKGYPWVWKDL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
Db 230 LDGFRIDAVKHKYSTFRDLWTHVRNATGKEMFAVAEFWKNLGALENYLNKTNWHS-- 287  
QY 241 AKVDFPALLYKDEAFDNKNI PALVSALONGOTVVSRRDPFKAVTFVANHTD----- 292  
Db 288 --VFDVPLHYNLYNA--SNSGGNYDMAKLLNG--TVQKHPHVAFTFVDNHDSPGESLESF 343  
QY 293 -IWNKPYAFAYFILTVE--GOPTIFYRDY-----EELWLNKDKLNLIWIHENLAGGSTDI 344  
Db 344 VQEWFKPLAYALLLTREQGYESVFGYGYGIPTHSVPAWMAKIDPILARQNFAYG-TQH 402  
QY 345 VYDNDLI-FVRNGYDGP--GLITYINLGSSKAGRWVYV-PKFAACIHEYTGNLGWS 400  
Db 403 DYFDHNIIGTREGNTHPNSGLATIMSDGPG--GEKMYVVGONKAGQVWHDITGNKPGT 461  
QY 401 VDKYVSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

## RESULT 2

US-08-861-837-2  
; Sequence 2, Application US/08861837  
; Patent No. 5856164  
; GENERAL INFORMATION:  
; APPLICANT: Ottup, Helle  
; APPLICANT: Bisgard-Frantzen, Henrik  
; APPLICANT: Ostergaard, Peter Rahbek  
; APPLICANT: Rasmussen, Michael Dolberg  
; APPLICANT: Van Der Zee, Pia  
; TITLE OF INVENTION: Alkaline Bacillus Amylase  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 58561640 No. 5856164disk of No. 5856164th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/861.837  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/446,803  
; FILING DATE: 01-June-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harrington, James J.  
; REGISTRATION NUMBER: 38,711  
; REFERENCE/DOCKET NUMBER: 4157.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 867-0123  
; TELEFAX: (212) 878-9655

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 485 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-861-837-2  
  
Query Match 19.8%; Score 488; DB 2; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;  
  
QY 11 GVIMQAFYNDVPSGGIWDITRQIKIPEDYDAGISAIWIPPAASKGMGGAYSGYDYDFD 70  
Db 7 GTWQYFPEWHLNDGNHNRRLDDASNLNRGITAIWIPAWKGT-SQNDVGYGAYDLYD 65  
QY 71 LGEDYDQKGTETRGSKQELVNMINTAHAYGMKVADIIVINHRAGGD-----LEWNP- 122  
Db 66 LGDFNOKGTETRTKGTGRSQAISALHAKNNGVQVYGVVNMHKGADATENVLAVEVNP 125  
QY 123 -----FVNDYT---WTFDSKVASKYTANYLDHFHNPHELHAGDGTGFGYPDICHDKSWDQ 174  
Db 126 NRNQISGDYITAEATKDFPGRGN---TYSDFKWRWYH-----FDGV-----DWDQ 169  
QY 175 -----YWLWASQES-----YAA-----YLRSIG 192  
Db 170 SRQFNRIYKFRGKAWDEVDSENGNYDYLMDYADVMDHPEVNVNLRWGEWYNTLN 229  
QY 193 IDAWRFDYVKGYPWVWKDL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
Db 230 LDGFRIDAVKHKYSTFRDLWTHVRNATGKEMFAVAEFWKNLGALENYLNKTNWHS-- 287  
QY 241 AKVDFPALLYKDEAFDNKNI PALVSALONGOTVVSRRDPFKAVTFVANHTD----- 292  
Db 288 --VFDVPLHYNLYNA--SNSGGNYDMAKLLNG--TVQKHPHVAFTFVDNHDSPGESLESF 343  
QY 293 -IWNKPYAFAYFILTVE--GOPTIFYRDY-----EELWLNKDKLNLIWIHENLAGGSTDI 344  
Db 344 VQEWFKPLAYALLLTREQGYESVFGYGYGIPTHSVPAWMAKIDPILARQNFAYG-TQH 402  
QY 345 VYDNDLI-FVRNGYDGP--GLITYINLGSSKAGRWVYV-PKFAACIHEYTGNLGWS 400  
Db 403 DYFDHNIIGTREGNTHPNSGLATIMSDGPG--GEKMYVVGONKAGQVWHDITGNKPGT 461  
QY 401 VDKYVSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

## RESULT 3

US-08-600-656-2  
; Sequence 2, Application US/08600656  
; Patent No. 6093562  
; GENERAL INFORMATION:  
; APPLICANT: Bisgard-Frantzen, Henrik  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; TITLE OF INVENTION: AMYLASE VARIANTS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 60935620 No. 6093562disk of No. 6093562th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/600,656  
FILING DATE: 13-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4318.204-US  
TELEPHONE: 212 867 0123  
TELEFAX: 212 867 0298  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 485 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-600-656-2

Query Match 19.8%; Score 488; DB 3; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.le-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;  
QY 11 GVIMQAFYVDVPSGGIWMDTIRQKIPWYDAGISAIWIPASKMGGAISMGYDYPDFD 70  
DB 7 GTMGOYFEWHLNDGHNRLRDDASNLNRGITAIPPAWKGT-SQNDVGYGAYDLID 65  
QY 71 LGEYDQKGTVETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
DB 66 LGEFNQKGTVRTKYGTRSQLESALHAKNGVQVYGVVNMHKGADATENVLAVENPN 125  
QY 123 -----FVNDYT---WTFDFKVASGKYTANYLDHPHNLHAGSGTGGYPDICHDKSWDQ 174  
DB 126 NRNQEISGDYTBIAWTKFDFPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169  
QY 175 -----YWLWASQES-----YAA-----YLSIG 192  
DB 170 SRQFQNRITYKFRGDKAWDEVDSENGNYDLYMADVMDHPEVNVNLRWGEWYNTLN 229  
QY 193 IDAMRFDYKGYAPVVKDWL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
DB 230 LDGFRIDAVKHIKYFTFRDLTHVRNATGEMFAVAFWKNLGDALLENLKNKTWNHS-- 287  
QY 241 AKVFDFALYKQDEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHTD----- 292  
DB 288 --VFDVPLHYNLNA-SNSGGNYDMAKLLNG-TVQKHPHMAVTFVNDHDSQPQGESLEF 343  
QY 293 -IIWNKYPAYAFILTYE-QOPTIFYRDI-----BEWLKDKLKNLIWIHENLAGGSTDI 344  
DB 344 VQEWFKPLAYALILTREQQPSVFGYDYGIPTHSVPAKAKIDPILARQNFAYG-TQH 402  
QY 345 VYDNDLI-FVRNGYDKP--GLITYINLGSSKAGRWYV-PKFAGACIHEITGNLGGW 400  
DB 403 DYFDHNIIGTWREGNTHPNSGLATIMSDGPG-GEKMWYVQGNKAGQVWHDTGNKPGT 461  
QY 401 VDKYVSSGWYLEAPAYDPANGQYGSVW 430  
DB 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 4  
US-09-170-670-2  
Sequence 2, Application US/09170670  
Patent No. 6187576  
GENERAL INFORMATION:  
APPLICANT: Svendsen, Allan  
APPLICANT: Borchert, Torben  
APPLICANT: Bisgard-Frantzen Henrik  
TITLE OF INVENTION: Alpha-Amylase Mutants  
FILE REFERENCE: 5276.200-US  
CURRENT APPLICATION NUMBER: US/09/170,670  
CURRENT FILING DATE: 1998-10-13  
EARLIER APPLICATION NUMBER: 1172/97

EARLIER FILING DATE: 1997-10-13  
EARLIER APPLICATION NUMBER: 60/063,306  
EARLIER FILING DATE: 1997-10-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 485  
TYPE: PRT  
ORGANISM: Bacillus sp.  
US-09-170-670-2  
Query Match 19.8%; Score 488; DB 3; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.le-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;  
QY 11 GVIMQAFYVDVPSGGIWMDTIRQKIPWYDAGISAIWIPASKMGGAISMGYDYPDFD 70  
DB 7 GTMGOYFEWHLNDGHNRLRDDASNLNRGITAIPPAWKGT-SQNDVGYGAYDLID 65  
QY 71 LGEYDQKGTVETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
DB 66 LGEFNQKGTVRTKYGTRSQLESALHAKNGVQVYGVVNMHKGADATENVLAVENPN 125  
QY 123 -----FVNDYT---WTFDFKVASGKYTANYLDHPHNLHAGSGTGGYPDICHDKSWDQ 174  
DB 126 NRNQEISGDYTBIAWTKFDFPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169  
QY 175 -----YWLWASQES-----YAA-----YLSIG 192  
DB 170 SRQFQNRITYKFRGDKAWDEVDSENGNYDLYMADVMDHPEVNVNLRWGEWYNTLN 229  
QY 193 IDAMRFDYKGYAPVVKDWL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
DB 230 LDGFRIDAVKHIKYFTFRDLTHVRNATGEMFAVAFWKNLGDALLENLKNKTWNHS-- 287  
QY 241 AKVFDFALYKQDEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHTD----- 292  
DB 288 --VFDVPLHYNLNA-SNSGGNYDMAKLLNG-TVQKHPHMAVTFVNDHDSQPQGESLEF 343  
QY 293 -IIWNKYPAYAFILTYE-QOPTIFYRDI-----BEWLKDKLKNLIWIHENLAGGSTDI 344  
DB 344 VQEWFKPLAYALILTREQQPSVFGYDYGIPTHSVPAKAKIDPILARQNFAYG-TQH 402  
QY 345 VYDNDLI-FVRNGYDKP--GLITYINLGSSKAGRWYV-PKFAGACIHEITGNLGGW 400  
DB 403 DYFDHNIIGTWREGNTHPNSGLATIMSDGPG-GEKMWYVQGNKAGQVWHDTGNKPGT 461  
QY 401 VDKYVSSGWYLEAPAYDPANGQYGSVW 430  
DB 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 5  
US-09-170-670-8  
Sequence 8, Application US/09170670  
Patent No. 6187576  
GENERAL INFORMATION:  
APPLICANT: Svendsen, Allan  
APPLICANT: Borchert, Torben  
APPLICANT: Bisgard-Frantzen Henrik  
TITLE OF INVENTION: Alpha-Amylase Mutants  
FILE REFERENCE: 5276.200-US  
CURRENT APPLICATION NUMBER: US/09/170,670  
CURRENT FILING DATE: 1998-10-13  
EARLIER APPLICATION NUMBER: 1172/97  
EARLIER FILING DATE: 1997-10-13  
EARLIER APPLICATION NUMBER: 60/063,306  
EARLIER FILING DATE: 1997-10-28  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 485  
TYPE: PRT

```

; ORGANISM: Bacillus sp.
US-09-170-670-8
Query Match      19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGAGYSGMYPDYPDFD 70
DB 7 GTMMQYFEWHLFNDGNHNRDRDASNLNRNGITAIWIPPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVEFRGSKQELVNMINTAHAYGMKVADIIVINHRAGD-----LEWNP- 122
DB 66 LGEFNQKGTVTRKYGTRSQLESIAIHALKKNQGVQVYGVVMMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTDFSKVASGKYTANYLDHPHNLHAGDSGTGGYPDICHDKSDQ 174
DB 126 NRNOEISGDYIEAWTKDFPPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRSIG 192
DB 170 SRQFQNRYYKFRGDKAWDNEVDSENGNYDLYMADVDMHDPVNVNELRWGEWYNTLN 229

QY 193 IDAWRFYVKGYPAPVWVKWL-----NWMGG---WAVEGYMDTNDVAV-----LNWYSSG 240
DB 230 LDGFRIDAVKHKYSFTRDMLTHVRNATGKEMFAVAEFKNDLGALENYLNKTNWHS-- 287

QY 241 AKVDFPALLYKMDAEAFDNKNIIPALVSALONGQTVSRDPFKAVTFVANHTD----- 292
DB 288 --VDFVPLHYNLYNA-SNSGNGYDMAKLLNG-TVQKHPMHAVTVDNHDSDQGESLESF 343

QY 293 -LIWKNYPAYAFILTYE-GOPTIFYRDY-----EELWKNCKLNLWIHENLAGGSTDI 344
DB 344 VQEWFKELAYALLITREQYPSVFYGYGIPHTSVPAKAKIDPILAEARQNFAYG-TQH 402

QY 345 VYDNDLELI-FVRNGYGDKP--GLITYINLGSSKAGRWVYV-PKFAGACIHEVTGNLGGW 400
DB 403 DYFDHNIIGWTREGNTHPNSGLATIMSDGPG-GEKMWYVGQNKAGQVWHHDITGNKPGT 461

QY 401 VDKVYSSGWLYLEAPAYDPANGQYGSVM 430
DB 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 6
US-09-193-068-2
; Sequence 2, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjruliff, S ren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-193-068-2

Query Match      19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGAGYSGMYPDYPDFD 70
DB 7 GTMMQYFEWHLFNDGNHNRDRDASNLNRNGITAIWIPPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVEFRGSKQELVNMINTAHAYGMKVADIIVINHRAGD-----LEWNP- 122
DB 66 LGEFNQKGTVTRKYGTRSQLESIAIHALKKNQGVQVYGVVMMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTDFSKVASGKYTANYLDHPHNLHAGDSGTGGYPDICHDKSDQ 174
DB 126 NRNOEISGDYIEAWTKDFPPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRSIG 192
DB 170 SRQFQNRYYKFRGDKAWDNEVDSENGNYDLYMADVDMHDPVNVNELRWGEWYNTLN 229

QY 193 IDAWRFYVKGYPAPVWVKWL-----NWMGG---WAVEGYMDTNDVAV-----LNWYSSG 240
DB 230 LDGFRIDAVKHKYSFTRDMLTHVRNATGKEMFAVAEFKNDLGALENYLNKTNWHS-- 287

QY 241 AKVDFPALLYKMDAEAFDNKNIIPALVSALONGQTVSRDPFKAVTFVANHTD----- 292
DB 288 --VDFVPLHYNLYNA-SNSGNGYDMAKLLNG-TVQKHPMHAVTVDNHDSDQGESLESF 343

QY 293 -LIWKNYPAYAFILTYE-GOPTIFYRDY-----EELWKNCKLNLWIHENLAGGSTDI 344
DB 344 VQEWFKELAYALLITREQYPSVFYGYGIPHTSVPAKAKIDPILAEARQNFAYG-TQH 402

QY 345 VYDNDLELI-FVRNGYGDKP--GLITYINLGSSKAGRWVYV-PKFAGACIHEVTGNLGGW 400
DB 403 DYFDHNIIGWTREGNTHPNSGLATIMSDGPG-GEKMWYVGQNKAGQVWHHDITGNKPGT 461

QY 401 VDKVYSSGWLYLEAPAYDPANGQYGSVM 430
DB 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 7
US-09-193-068-8
; Sequence 8, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjruliff, S ren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-193-068-8

Query Match      19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPAKMGAGYSGMYPDYPDFD 70
DB 7 GTMMQYFEWHLFNDGNHNRDRDASNLNRNGITAIWIPPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVEFRGSKQELVNMINTAHAYGMKVADIIVINHRAGD-----LEWNP- 122
DB 66 LGEFNQKGTVTRKYGTRSQLESIAIHALKKNQGVQVYGVVMMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTDFSKVASGKYTANYLDHPHNLHAGDSGTGGYPDICHDKSDQ 174
DB 126 NRNOEISGDYIEAWTKDFPPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRSIG 192
DB 170 SRQFQNRYYKFRGDKAWDNEVDSENGNYDLYMADVDMHDPVNVNELRWGEWYNTLN 229

```

Qy	193	IDAWREDDYKGYAPVVKDWL-----NWTGG--WAVEGYTNDTNDVAV-----LNWAYSSG 240
Db	230	LDGFRIDAVKHKIYGFTRDDLUHVRNATGKEMFAVAFKWNLDGALENLYNKTNWNSH-----287
Qy	241	AKYDFDALYKMDAEAFDNKNIPALVASALQNGQTVVSRDPPKAVTFVANHDTD-----292
Db	288	--VFDVPLHYNLYNA-SNSGGNYDMAKLLNG-TVVQKEPMAVTFVVDNHDHSDQGESLESF 343
Qy	293	-IWNKYPAYAFILIYE-GQTFIFRDY-----EEWLNKDKUKULIWHENLAGGSTDI 344
Db	344	VOEWFKPLAYAILITREQQYPSVFGDYGYPTHSVPAMKAKIDPILFARQNPAYG-TQH 402
Qy	345	VYVDNDELT-FVRNGYGDKP--GLITYINLGSSKAGRWVVV-PKFAGACIHEVTGNLGW 400
Db	403	DYFDHNNIIGWTRFEGNTTHPSNGLATIMSDGP-GEKWMYVGNQKAGQVWHDTGNKPKET 461
Qy	401	VDKYIYSSGWYYLEAPYDPAWGQYGSVW 430
Db	462	VT--INADGW-----ANFVSNGG-SYSIW 482

```

RESULT 8
US-09-183-412-2
; Sequence 2, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-183-412-2

```

[illegible]

Qy	193	IDAWREDVYKGYAPVVDKTL---	NWGG--	MAVEYWDNTVDV-----	LNWAYSSG	240
Db	230	LOGFRIDAVKHIKYGFTRDTHV	RNATGKEMFAVEFWKNDJG	ALLENLYNKTNWNHS---	287	
Qy	241	AKVFDFALYKMDFAFDKNKIPAL	VSAALQNSQTWVSDDPKAVTF	VANHDTD-----	292	
Db	288	--VDFVPLHYNLNA--SNSGGNYD	MAKLLNG-TVQKHPMHAVTF	VDNHDSDQSGESLESF	343	
Qy	293	-IWNKYPAYAFILTYE--GQTI	IFYRDY-----	EWLNKDKUKULIWHENI	AGSSTDI	344
Db	344	VOEFKPLAYAILITREGQPSV	GYDYGYPITHSVFAMKAKITD	PILFARQNFAYG--TOH	402	
Qy	345	VYVDNDELI--FVRNRYGSKP--	GLITVINLGSSKAGRWVV--	PKFAGACIHEVTGNLGGW	400	
Db	403	DYFDHNNIIGTREGNTTHPSN	GLATTMSDGP-GEKWMYVGQ	NKAGQVWHIDITGNKPGT	461	
Qy	401	VDKYVYSSGWVYLEAPYDPA	NGQYGSVW	430		
Db	462	VT--INADGW-----ANFS	VNGG--SVSIIW	482		

```

RESULT 9
US-09-183-412-8
; Sequence 8, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjørne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kiaerulff, Søren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-183-412-8

```

Query Match	19.8%;	Score 488;	DB 3;	Length 485;
Best Local Similarity	31.0%;	Fred. No. 1.1e-34;		
Matches 158;	Conservative	59;	Mismatches 169;	Indels 124; Gaps 27;
QY	11	GVIMQAFYDWPVSGGIWMDTIIRKQIPEWYDAGISAIMIWPASKMGGAYSMDGYDFFD	70	
Db	7	GTMMQYFEWHLPLNDGNHWRLLDDASLNLRGCIITAIWIPPAWKGT-SQNDYGVGAYDLVD	65	
QY	71	LGEDYDQGTGYETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGD-----LENNP-	122	
Db	66	LGFEHQGTGRTKYGRFSQLESAIHALKXGVQVYGVDMNHKGADATENVLAVEVNP	135	
QY	123	-----FVNDY---WTFDFSKVASGKYTANYIDFHNELHAGDSGTFGGYDPICHDKSDWQ	174	
Db	126	NRNQETSGDYITAEAWTKDFDPGRGN---TYSDFKRWYH---FDGV-----DWDQ	169	
QY	175	-----YWLWASQES-----YAA-----YLRSG-	192	
Db	170	SRQFQNRKYKFRGDGKAWDWEVDSENGNYDYLMTADVMDHDPVVNLRMRGGEWYNTLN	229	
QY	193	IDAWRFDYKGAIPWYVKDWL-----NWKGG--WAVGEYWDINVDV-----LNWAYSSG	240	

Db 230 LDGFRIDAVKHKVSFTFRDMLTHVRNATGKEMFAVAEFKXNDLGALENYLNKTNWHS-- 287  
Qy 241 AKVDFDALYKMDAFAFNKNIPALVSALQNGQTVSRDPFKAVTFVANHDTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGGNYDMAKLNG-TVQKHPHVAFTFVDNHDSPQGSLSF 343  
Qy 293 -LIWKNYPAYAFILTYE-GQTFIYRDY-----EWLNDKDKLNLIWHENLAGSTDI 344  
Db 344 VQEWKPLAYALILTRREGYPSVFGYGYGIPHTSVFAMKAKDPILEARQNFAYG-TQH 402  
Qy 345 VYDNDLI-FVRNGYDKP--GLITYINLSSKAGRWVYV-PKFAGACIHEYTGNLGWM 400  
Db 403 DYFDHNIIGWTRGNTHPNGLATIMSDPG-GEKWMYVGONKAGQVWHDTGNKPGT 461  
Qy 401 VDKYVYSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 10  
US-09-264-097-5  
; Sequence 5, Application US/09264097  
; Patent No. 6287826  
; GENERAL INFORMATION:  
; APPLICANT: No. 6287826man, Barrie Edmund  
; APPLICANT: Hendriksen, Hanne Vang  
; TITLE OF INVENTION: Enzymatic Preparation of Glucose Syrup  
; TITLE OF INVENTION: From Starch  
; FILE REFERENCE: 5278-200-US  
; CURRENT APPLICATION NUMBER: US/09/264,097  
; EARLIER FILING DATE: 1999-03-08  
; EARLIER FILING DATE: 1998-03-09  
; EARLIER APPLICATION NUMBER: 60/079,209  
; EARLIER FILING DATE: 1998-03-24  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus  
US-09-264-097-5

Query Match  
Best Local Similarity 19.8%; Score 488; DB 3; Length 485;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

Qy 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGYDYPDFD 70  
Db 7 GTMMQYFEWHLFNDGNHNRRLDASNLNRGITAIWIPPAWKT-SQNDVGYGAYDLYD 65  
Qy 71 LGEYDQKGTVTREFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVTKGTGRSQALESALHAKNGVQVYGVVWVNHKGGADATENVLAVENPN 125  
Qy 123 -----FVNDYT---WTDFSKVASGYTANYLDLPHNELHAGDSGTGFGYDICHDKSWDQ 174  
Db 126 NRNOEISGDYTIETATKDFPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169  
Qy 175 -----YWLWASQES-----YAA-----YLRSTG 192  
Db 170 SRQFNRIYKPRGDKADWDEVSNGNYDLYADVDMHDPVNVNELRWGEWNTINL 229  
Qy 193 IDAWRFYKGYAPWVYKDWL-----NMWGG--WAVEGYDNTVDV-----LNWYSSG 240  
Db 230 LDGFRIDAVKHKVSFTFRDMLTHVRNATGKEMFAVAEFKXNDLGALENYLNKTNWHS-- 287  
Qy 241 AKVDFDALYKMDAFAFNKNIPALVSALQNGQTVSRDPFKAVTFVANHDTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGGNYDMAKLNG-TVQKHPHVAFTFVDNHDSPQGSLSF 343  
Qy 293 -LIWKNYPAYAFILTYE-GQTFIYRDY-----EWLNDKDKLNLIWHENLAGSTDI 344

Db 344 VQEWKPLAYALILTRREGYPSVFGYGYGIPHTSVFAMKAKDPILEARQNFAYG-TQH 402  
Qy 345 VYDNDLI-FVRNGYDKP--GLITYINLSSKAGRWVYV-PKFAGACIHEYTGNLGWM 400  
Db 403 DYFDHNIIGWTRGNTHPNGLATIMSDPG-GEKWMYVGONKAGQVWHDTGNKPGT 461  
Qy 401 VDKYVYSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 11  
US-09-354-191A-2  
; Sequence 2, Application US/09354191A  
; Patent No. 6297038  
; GENERAL INFORMATION:  
; APPLICANT: Bisgard-Frantzen, Henrik  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; TITLE OF INVENTION: AMYLASE VARIANTS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6297038o No. 6297038disk of No. 6297038th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/354,191A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/600,656  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4318.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 867 0123  
; TELEFAX: 212 867 0298  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 485 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
US-09-354-191A-2

Query Match  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

Qy 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGYDYPDFD 70  
Db 7 GTMMQYFEWHLFNDGNHNRRLDASNLNRGITAIWIPPAWKT-SQNDVGYGAYDLYD 65  
Qy 71 LGEYDQKGTVTREFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVTKGTGRSQALESALHAKNGVQVYGVVWVNHKGGADATENVLAVENPN 125  
Qy 123 -----FVNDYT---WTDFSKVASGYTANYLDLPHNELHAGDSGTGFGYDICHDKSWDQ 174  
Db 126 NRNOEISGDYTIETATKDFPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169

QY 175 -----YMLWASQES-----YAA-----YLSIG 192  
DB 170 SRQFNRIYKFRGDKAMDWEVDSENGNYDYLMDYADVDMDHPEVNVNELRRGWETNLN 229  
QY 193 IDAWREFDVVKGAPVWVWDL-----NWGG--WAVEGYMDTNDVAV-----LNWAYSSG 240  
DB 230 LDGFIDAVKHKISFTDLWTHVRNATGKEMFAVAEFWKNDLGALENLKNKTWNHS-- 287  
QY 241 AKVFDPALYYKMDAFAFNKKNIPALVSALQNGQTVVSRDPFKAVTFVANHTD----- 292  
DB 288 --VFDVPLHYNLYNA--SNSGGNYDMAKLLNG--TVQKHPEHVAIFVDNHDQPGESLESF 343  
QY 293 --LIWKNKYAPAYAFILTYE--GQPTIFRDIY-----EELNKKLNIWIHENLAGGSDTI 344  
DB 344 VOEWFKPLAYALILIREQGYPSVFYGYGIPTHSVPAKAKIDPILFARQNFAYG--TQH 402  
QY 345 VYDNDDELII-FVRNGYDKP--GLITYINLGSSKAGRWVYV--PKFAGACIHEYTGNLGGW 400  
DB 403 DYFDHNNIIGWTREGNTTHPNSGLATIMSDGPG--GEKWMYVGQKAGQVWHDITGNKPGT 461  
QY 401 VDKYVYSSGWVLEAPAYDPDPANGQYGSVM 430  
DB 462 VT--INADGW-----ANFSVNGG--SVSIV 482

RESULT 12  
US-09-291-023A-4  
; Sequence 4, Application US/09291023A  
; Patent No. 6309871  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic A  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/291,023A  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: DK 1999 00438  
; PRIOR FILING DATE: 1999-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-291-023A-4

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 30.5%; Pred. No. 1.1e-34;  
Matches 154; Conservative 68; Mismatches 169; Indels 114; Gaps 24;

QY 11 GVIMQAFYWDVPSSGIWMDTIRQKIPWYDAGISAIWIPIPPASKMGAYSGMDPYDPDFD 70  
DB 7 GTIMQYFEWNVDPDGOHNRHNNNAQNKNAGITAIWIPIPAWKT--SQNDVGYGAYDLYD 65  
QY 71 LGEYDQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP 123  
DB 66 LGEFNQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP 125  
QY 124 VNDY-----TWTFPSKVASGKYTANY-----LDFHPNELHAGDSGTFGGYPD 165  
DB 126 NRNQEVSGTYQIEAWTGFNPGCGNHSSFKRWYHFDGTDQSRQLANRIYKFRG--- 182  
QY 166 ICHDKSWD-----QYWLWASQES-----YAYLRSIGIDAWRFDYVKG 203  
DB 183 --DGRKAMDWEVDTEGNYDYLMDYADVDMDHPEVINELNRGWVYANTLNLDGFRDLAVKH 240  
QY 204 YAPVWVKDNLNWGG-----WAVEGYMDTNDVAV-----LNWAYSSGAKVDFEALYYK 251  
DB 241 IKFSFMRDLGHVRGQTKNLFVAEYWKNDLGALENLKNKTWNHSA-----FDVPLHYN 296

QY 252 MDEA-----FDKNKIPALVSALQNGQTVVSRDPFKAVTFVANHTD-----LIWN 296  
DB 297 LYQASNSNGYDMRN-----LLNG--TLVQRHPSHAVTFVDNHDTPQGEALESFVQGW 348  
QY 297 KYPAAYAFILTYE--GQPTIFRDIYEEWLN-----KDKLKNLIWIHENLAGGSDTIYVDN 349  
DB 349 KPLAYATILIREQGYQVYGYGIPSDGVPYSYRQOIDPILLKARQQYAYGRQH--DYFDH 407  
QY 350 DELI--FVRNGYDKP--GLITYINLGSSKAGRWVYV--PKFAGACIHEYTGNLGGWVYV 405  
DB 408 WDWIGWTREGNASHPNSGLATIMSDGPG--GSKWMYVGROKAGEVWHDITGNRSGITV--I 464  
QY 406 YSSGWVLEAPAYDPDPANGQYGSVM 430  
DB 465 NQDQGWGHEF-----VNGG--SVSIV 482

RESULT 13  
US-09-291-023A-14  
; Sequence 14, Application US/09291023A  
; Patent No. 6309871  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic A  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/291,023A  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: DK 1999 00438  
; PRIOR FILING DATE: 1999-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus  
US-09-291-023A-14

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 30.5%; Pred. No. 1.1e-34;  
Matches 154; Conservative 68; Mismatches 169; Indels 114; Gaps 24;

QY 11 GVIMQAFYWDVPSSGIWMDTIRQKIPWYDAGISAIWIPIPPASKMGAYSGMDPYDPDFD 70  
DB 7 GTIMQYFEWNVDPDGOHNRHNNNAQNKNAGITAIWIPIPAWKT--SQNDVGYGAYDLYD 65  
QY 71 LGEYDQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP 123  
DB 66 LGEFNQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP 125  
QY 124 VNDY-----TWTFPSKVASGKYTANY-----LDFHPNELHAGDSGTFGGYPD 165  
DB 126 NRNQEVSGTYQIEAWTGFNPGCGNHSSFKRWYHFDGTDQSRQLANRIYKFRG--- 182  
QY 166 ICHDKSWD-----QYWLWASQES-----YAYLRSIGIDAWRFDYVKG 203  
DB 183 --DGRKAMDWEVDTEGNYDYLMDYADVDMDHPEVINELNRGWVYANTLNLDGFRDLAVKH 240  
QY 204 YAPVWVKDNLNWGG-----WAVEGYMDTNDVAV-----LNWAYSSGAKVDFEALYYK 251  
DB 241 IKFSFMRDLGHVRGQTKNLFVAEYWKNDLGALENLKNKTWNHSA-----FDVPLHYN 296  
QY 252 MDEA-----FDKNKIPALVSALQNGQTVVSRDPFKAVTFVANHTD-----LIWN 296  
DB 297 LYQASNSNGYDMRN-----LLNG--TLVQRHPSHAVTFVDNHDTPQGEALESFVQGW 348  
QY 297 KYPAAYAFILTYE--GQPTIFRDIYEEWLN-----KDKLKNLIWIHENLAGGSDTIYVDN 349

Db 349 KPLATATILTRQGYQVYFYGYGIPSDGVPYSRQIDPLLKARQYAYGRQH-DYFDH 407  
QY 350 DELI-FVRNGYDGP--GLITVINIGSSKAGRWVTPK-FAGACIHEYTNLGNWYKYV 405  
Db 408 MDVIGWTRGNASHPNGLATIMSDPG-GSKWMTVGRKAGVWHDMTGNRSGTGT--I 464  
QY 406 YSSGKYLEAPAYDPANGQYGSVW 430  
Db 465 NODMGHFF-----VNGG-SVSVW 482

RESULT 14  
US-09-290-734-2  
; Sequence 2, Application US/09290734  
; Patent No. 6361989  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; APPLICANT: Bisgard-Frantzen Henrik  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Nielsen, Bjarne Ronfeldt  
; APPLICANT: Nielsen, Vibeke Skovgaard  
; APPLICANT: Hoeck, Lisbeth Hedegaard  
; TITLE OF INVENTION: No. 6361989el -Amylase And -Amylase Mutants  
; FILE REFERENCE: 5276.400-US  
; CURRENT APPLICATION NUMBER: US/09/290,734  
; CURRENT FILING DATE: 1999-04-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-290-734-2

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYMDVPSGGIWMDTIRQKIPWYDAGISAIWIIPPASKMGAGYSGMYPDYDFD 70  
Db 7 GTMWQYFEWHLNDGNHNRRLDASNLNRGITAIIWIPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVETRFPGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGVQVGYGVVNMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTFDSKVASGKYTANYLDHPNELHAGDSGTGGYPDICHDKSWDQ 174  
Db 126 NRNOEISGDYITIEAWTKFDPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRISG 192  
Db 170 SRQFONRIYKFRGDKAWDNEVDSENGNYDLYADVDMDPHPEVNMELRWGEWYNTLN 229

QY 193 IDAWRPDYKGYAPVWVKDWL---NWGG--WAVEGYWDTNVDAV-----LNWAYSSG 240  
Db 230 LDGFRIDAVKHIIKYSFTRDWTHTVRNATGKEMFAVAFWKNLGALENYLNKTNWHS-- 287

QY 241 AKVDFDALYKMDAEAFDNKNIIPALVSALONGQTVSRDPFKAVTFVANHTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGNYDMAKLNG--TVQKHPHMAVTFVDNHDSPGESLESF 343

QY 293 -IWNKYPAYAFILTYE-GOTTIFYRDI-----EWLKDKLKNLIWIHENLAGSTDI 344  
Db 344 VQEWFKPLAYALILITREQGYPVFGYGYGIPTHSVPMKAKIDPILFARQNFAYG-TQH 402

QY 345 VYDNDLI-FVRNGYDGP--GLITVINIGSSKAGRWVTPK-FAGACIHEYTNLGNW 400  
Db 403 DYFDHNIIGWTRGNTHPNGLATIMSDPG-GEKWMYVGNKAGQVWHDITGNKPGT 461

QY 401 VDKVYSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSVW 482

Db 462 VT--INADGW-----ANFSVNGG-SVSVW 482

RESULT 15  
US-09-290-734-8  
; Sequence 8, Application US/09290734  
; Patent No. 6361989  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; APPLICANT: Bisgard-Frantzen Henrik  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Nielsen, Bjarne Ronfeldt  
; APPLICANT: Nielsen, Vibeke Skovgaard  
; APPLICANT: Hoeck, Lisbeth Hedegaard  
; TITLE OF INVENTION: No. 6361989el -Amylase And -Amylase Mutants  
; FILE REFERENCE: 5276.400-US  
; CURRENT APPLICATION NUMBER: US/09/290,734  
; CURRENT FILING DATE: 1999-04-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-290-734-8

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYMDVPSGGIWMDTIRQKIPWYDAGISAIWIIPPASKMGAGYSGMYPDYDFD 70  
Db 7 GTMWQYFEWHLNDGNHNRRLDASNLNRGITAIIWIPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVETRFPGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGVQVGYGVVNMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTFDSKVASGKYTANYLDHPNELHAGDSGTGGYPDICHDKSWDQ 174  
Db 126 NRNOEISGDYITIEAWTKFDPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRISG 192  
Db 170 SRQFONRIYKFRGDKAWDNEVDSENGNYDLYADVDMDPHPEVNMELRWGEWYNTLN 229

QY 193 IDAWRPDYKGYAPVWVKDWL---NWGG--WAVEGYWDTNVDAV-----LNWAYSSG 240  
Db 230 LDGFRIDAVKHIIKYSFTRDWTHTVRNATGKEMFAVAFWKNLGALENYLNKTNWHS-- 287

QY 241 AKVDFDALYKMDAEAFDNKNIIPALVSALONGQTVSRDPFKAVTFVANHTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGNYDMAKLNG--TVQKHPHMAVTFVDNHDSPGESLESF 343

QY 293 -IWNKYPAYAFILTYE-GOTTIFYRDI-----EWLKDKLKNLIWIHENLAGSTDI 344  
Db 344 VQEWFKPLAYALILITREQGYPVFGYGYGIPTHSVPMKAKIDPILFARQNFAYG-TQH 402

QY 345 VYDNDLI-FVRNGYDGP--GLITVINIGSSKAGRWVTPK-FAGACIHEYTNLGNW 400  
Db 403 DYFDHNIIGWTRGNTHPNGLATIMSDPG-GEKWMYVGNKAGQVWHDITGNKPGT 461

QY 401 VDKVYSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSVW 482



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:35:47 ; Search time 49 Seconds  
(without alignments)  
2515.500 Million cell updates/sec

Title: US-10-081-739A-2  
Perfect score: 2459  
Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGGYGVSWVSCGVG 436

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2459	100.0	436	14	US-10-146-662-2
2	2459	100.0	436	14	US-10-105-733-2
3	2459	100.0	436	14	US-10-081-739A-2
4	2424	98.6	436	14	US-10-081-872-72
5	2424	98.6	436	14	US-10-385-305-16
6	2406	97.8	436	14	US-10-081-872-6
7	2406	97.8	436	14	US-10-385-305-6
8	2378	96.7	436	14	US-10-081-872-60
9	2378	96.7	436	14	US-10-385-305-60
10	2377	96.7	436	14	US-10-081-872-38
11	2377	96.7	436	14	US-10-385-305-38
12	2371	96.4	436	14	US-10-081-872-30
13	2371	96.4	436	14	US-10-385-305-30
14	2368	96.3	436	14	US-10-081-872-4
15	2368	96.3	436	15	US-10-385-305-4

16	2360	96.0	436	14	US-10-081-872-40	Sequence 40, Appl
17	2360	96.0	436	15	US-10-385-305-40	Sequence 40, Appl
18	2358	95.9	436	14	US-10-081-872-16	Sequence 16, Appl
19	2358	95.9	436	14	US-10-081-872-62	Sequence 62, Appl
20	2358	95.9	436	15	US-10-385-305-16	Sequence 16, Appl
21	2358	95.9	436	15	US-10-385-305-62	Sequence 62, Appl
22	2356	95.8	436	14	US-10-081-872-20	Sequence 20, Appl
23	2356	95.8	436	14	US-10-081-872-32	Sequence 32, Appl
24	2356	95.8	436	15	US-10-385-305-20	Sequence 20, Appl
25	2356	95.8	436	15	US-10-385-305-32	Sequence 32, Appl
26	2352	95.6	436	14	US-10-081-872-56	Sequence 56, Appl
27	2352	95.6	436	15	US-10-385-305-56	Sequence 56, Appl
28	2350	95.6	436	14	US-10-081-872-10	Sequence 10, Appl
29	2350	95.6	436	14	US-10-081-872-28	Sequence 28, Appl
30	2350	95.6	436	15	US-10-385-305-10	Sequence 10, Appl
31	2350	95.6	436	15	US-10-385-305-28	Sequence 28, Appl
32	2349	95.5	436	14	US-10-081-872-42	Sequence 42, Appl
33	2349	95.5	436	14	US-10-081-872-46	Sequence 46, Appl
34	2349	95.5	436	14	US-10-228-063-1	Sequence 1, Appl
35	2349	95.5	436	15	US-10-385-305-42	Sequence 42, Appl
36	2349	95.5	436	15	US-10-385-305-46	Sequence 46, Appl
37	2349	95.5	518	14	US-10-228-063-15	Sequence 15, Appl
38	2349	95.5	741	14	US-10-228-063-10	Sequence 10, Appl
39	2349	95.5	820	14	US-10-228-063-16	Sequence 16, Appl
40	2347	95.4	436	14	US-10-081-872-14	Sequence 14, Appl
41	2347	95.4	436	15	US-10-385-305-14	Sequence 14, Appl
42	2344	95.3	454	14	US-10-228-063-13	Sequence 13, Appl
43	2344	95.3	460	14	US-10-228-063-14	Sequence 14, Appl
44	2344	95.3	460	14	US-10-228-063-33	Sequence 33, Appl
45	2344	95.3	460	14	US-10-228-063-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1  
US-10-146-662-2  
; Sequence 2, Application US/10146662  
; Publication No. US20030013172A1  
; GENERAL INFORMATION:  
; APPLICANT: Grendash, Joel  
; TITLE OF INVENTION: NOVEL METHODS OF ENZYME PURIFICATION  
; FILE REFERENCE: 09010-109001  
; CURRENT APPLICATION NUMBER: US/10/146,662  
; CURRENT FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; TYPE: PRT  
; LENGTH: 436  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated polypeptide  
US-10-146-662-2

Query Match	100.0%	Score	2459;	DB	14;	Length	436;
Best Local Similarity	100.0%	Pred. No.	1.8e-219;				
Matches	436;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIRQKIPEWYDAGISAIWIWPPASKMGGGAYS	60				
Db	1	MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIRQKIPEWYDAGISAIWIWPPASKMGGGAYS	60				
QY	61	MGYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVITADIVINHRAGDLEW	120				
Db	61	MGYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVITADIVINHRAGDLEW	120				
QY	121	NPVVDYTWTFPSKVASKYTYNYLDLPHNELHAGDSGTFGYDPDICHDKSWDQYWLWAS	180				
Db	121	NPVVDYTWTFPSKVASKYTYNYLDLPHNELHAGDSGTFGYDPDICHDKSWDQYWLWAS	180				

us-10-081-739a-2.rapb

Wed Jun 30 10:27:02 2004

```

QY 181 QESYAAYLRSIGIDAWRFDYVKGYPVWVVDLNNWGGWAVGEYWDNTVDVAVLNWAYSSG 240
DB 181 QESYAAYLRSIGIDAWRFDYVKGYPVWVVDLNNWGGWAVGEYWDNTVDVAVLNWAYSSG 240
QY 241 AKVDFPALYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
DB 241 AKVDFPALYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
QY 301 YAFILTYEQPTIFYRDYEEWLNKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
DB 301 YAFILTYEQPTIFYRDYEEWLNKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
QY 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
DB 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
QY 421 ANGQYGSVWSYCGVG 436
DB 421 ANGQYGSVWSYCGVG 436

```

```

RESULT 2
US-10-105-733-2
; Sequence 2, Application US/10105733
; Publication No. US20030138786A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Miller, Carl
; APPLICANT: Kazaoka, Martin
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-107002
; CURRENT APPLICATION NUMBER: US/10105,733
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/081,739
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically engineered
US-10-105-733-2
Query Match 100.0%; Score 2459; DB 14; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.8e-219;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAKYSELEKGVINQAFYWDVPSGGIWDTRIROKIPEDWDAGISAIWIPSPKMGMGAYS 60
DB 1 MAKYSELEKGVINQAFYWDVPSGGIWDTRIROKIPEDWDAGISAIWIPSPKMGMGAYS 60
QY 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
DB 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
QY 121 NPFVNDYTWTFDSKVASKYKTANYLDHFNELHAGSGTGGYDPICHDKSWDQYWLWAS 180
DB 121 NPFVNDYTWTFDSKVASKYKTANYLDHFNELHAGSGTGGYDPICHDKSWDQYWLWAS 180
QY 181 QESYAAYLRSIGIDAWRFDYVKGYPVWVVDLNNWGGWAVGEYWDNTVDVAVLNWAYSSG 240

```

```

DB 181 QESYAAYLRSIGIDAWRFDYVKGYPVWVVDLNNWGGWAVGEYWDNTVDVAVLNWAYSSG 240
QY 241 AKVDFPALYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
DB 241 AKVDFPALYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
QY 301 YAFILTYEQPTIFYRDYEEWLNKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
DB 301 YAFILTYEQPTIFYRDYEEWLNKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
QY 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
DB 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
QY 421 ANGQYGSVWSYCGVG 436
DB 421 ANGQYGSVWSYCGVG 436

```

```

RESULT 3
US-10-081-739A-2
; Sequence 2, Application US/10081739A
; Publication No. US20030170634A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Miller, Carl
; APPLICANT: Kazaoka, Martin
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-107001
; CURRENT APPLICATION NUMBER: US/10/081,739A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically engineered
US-10-081-739A-2
Query Match 100.0%; Score 2459; DB 14; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.8e-219;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAKYSELEKGVINQAFYWDVPSGGIWDTRIROKIPEDWDAGISAIWIPSPKMGMGAYS 60
DB 1 MAKYSELEKGVINQAFYWDVPSGGIWDTRIROKIPEDWDAGISAIWIPSPKMGMGAYS 60
QY 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
DB 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
QY 121 NPFVNDYTWTFDSKVASKYKTANYLDHFNELHAGSGTGGYDPICHDKSWDQYWLWAS 180
DB 121 NPFVNDYTWTFDSKVASKYKTANYLDHFNELHAGSGTGGYDPICHDKSWDQYWLWAS 180
QY 181 QESYAAYLRSIGIDAWRFDYVKGYPVWVVDLNNWGGWAVGEYWDNTVDVAVLNWAYSSG 240
DB 181 QESYAAYLRSIGIDAWRFDYVKGYPVWVVDLNNWGGWAVGEYWDNTVDVAVLNWAYSSG 240

```

QY 241 AKVFDFALYYKMDFAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
| | | | |  
Db 241 AKVFDFALYYKMDFAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
| | | | |  
QY 301 YAFILTYEGQPTIFRYDYEEWLNKDKNLIIWHENLAGGSDTIYYDNDDELIFVRNGYG 360  
| | | | |  
Db 301 YAFILTYEGQPTIFRYDYEEWLNKDKNLIIWHENLAGGSDTIYYDNDDELIFVRNGYG 360  
| | | | |  
QY 361 DKPGLITYINLGSSKAGRWVVPKFGACIHEYTGNLGGWVDKVVYSSGWWYLEAPAYDP 420  
| | | | |  
Db 361 DKPGLITYINLGSSKAGRWVVPKFGACIHEYTGNLGGWVDKVVYSSGWWYLEAPAYDP 420  
| | | | |  
QY 421 ANGOYGYSWMSYCGVG 436  
| | | | |  
Db 421 ANGOYGYSWMSYCGVG 436  
| | | | |

## RESULT 4

US-10-081-872-72

; Sequence 72, Application US/10081872

; Publication No. US20030125534A1

; GENERAL INFORMATION:

; APPLICANT: Callen, Walter

; APPLICANT: Richardson, Toby

; APPLICANT: Frey, Gerhard

; APPLICANT: Short, Jay M.

; APPLICANT: Mathur, Eric J.

; APPLICANT: Gray, Kevin A.

; APPLICANT: Kerovuo, Janne S.

; APPLICANT: Slupska, Malgorzata

; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY

; FILE REFERENCE: 09010-108001

; CURRENT APPLICATION NUMBER: US/10/081,872

; PRIOR FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: US 60/270,495

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: US 60/270,496

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: US 60/291,122

; PRIOR FILING DATE: 2001-05-14

; NUMBER OF SEQ ID NOS: 321

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 72

; LENGTH: 436

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetically generated polypeptide

US-10-081-872-72

Query Match 98.6%; Score 2424; DB 14; Length 436;  
Best Local Similarity 98.4%; Pred. No. 3.2e-216;  
Matches 429; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWMWDTIRQKIPEDWDAGISAIWIIPASKMGGAYS 60  
| | | | |  
Db 1 MAKYLEEGGVIMQAFYWDVPSGGIWMWDTIRQKIPEDWDAGISAIWIIPASKMGGAYS 60  
| | | | |  
QY 61 MGYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120  
| | | | |  
Db 61 MGYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120  
| | | | |  
QY 121 NPFVNDYTWTFPSKVASKYKTANYLDHFNELHAGDSGTFGGYPDI CHDKSWDQYWLWAS 180  
| | | | |  
Db 121 NPFVNDYTWTFPSKVASKYKTANYLDHFNELHAGDSGTFGGYPDI CHDKSWDQYWLWAS 180  
| | | | |  
QY 181 QESYAAALRSIGIDAWRFDYVKGYAPWVVKDNLNMGWGWAVGEYWDTNVDVAVLNWAYSSG 240  
| | | | |  
Db 181 QESYAAALRSIGIDAWRFDYVKGYAPWVVKDNLNMGWGWAVGEYWDTNVDVAVLNWAYSSG 240  
| | | | |  
QY 241 AKVFDFALYYKMDFAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
| | | | |

Db 241 AKVFDFALYYKMDFAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
| | | | |  
QY 301 YAFILTYEGQPTIFRYDYEEWLNKDKNLIIWHENLAGGSDTIYYDNDDELIFVRNGYG 360  
| | | | |  
Db 301 YAFILTYEGQPTIFRYDYEEWLNKDKNLIIWHENLAGGSDTIYYDNDDELIFVRNGYG 360  
| | | | |  
QY 361 DKPGLITYINLGSSKAGRWVVPKFGACIHEYTGNLGGWVDKVVYSSGWWYLEAPAYDP 420  
| | | | |  
Db 361 DKPGLITYINLGSSKAGRWVVPKFGACIHEYTGNLGGWVDKVVYSSGWWYLEAPAYDP 420  
| | | | |  
QY 421 ANGOYGYSWMSYCGVG 436  
| | | | |  
Db 421 ANGOYGYSWMSYCGVG 436  
| | | | |

## RESULT 5

US-10-385-305-72

; Sequence 72, Application US/10385305

; Publication No. US20040018607A1

; GENERAL INFORMATION:

; APPLICANT: Callen, Walter

; APPLICANT: Richardson, Toby

; APPLICANT: Frey, Gerhard

; APPLICANT: Short, Jay M.

; APPLICANT: Mathur, Eric J.

; APPLICANT: Gray, Kevin A.

; APPLICANT: Kerovuo, Janne S.

; APPLICANT: Slupska, Malgorzata

; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY

; FILE REFERENCE: 09010-108001

; CURRENT APPLICATION NUMBER: US/10/385,305

; PRIOR FILING DATE: 2003-03-06

; PRIOR APPLICATION NUMBER: US/10/081,872

; PRIOR FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: US 60/270,495

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: US 60/270,496

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: US 60/291,122

; PRIOR FILING DATE: 2001-05-14

; NUMBER OF SEQ ID NOS: 321

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 72

; LENGTH: 436

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetically generated polypeptide

US-10-385-305-72

Query Match 98.6%; Score 2424; DB 15; Length 436;  
Best Local Similarity 98.4%; Pred. No. 3.2e-216;  
Matches 429; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWMWDTIRQKIPEDWDAGISAIWIIPASKMGGAYS 60  
| | | | |  
Db 1 MAKYLEEGGVIMQAFYWDVPSGGIWMWDTIRQKIPEDWDAGISAIWIIPASKMGGAYS 60  
| | | | |  
QY 61 MGYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120  
| | | | |  
Db 61 MGYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120  
| | | | |  
QY 121 NPFVNDYTWTFPSKVASKYKTANYLDHFNELHAGDSGTFGGYPDI CHDKSWDQYWLWAS 180  
| | | | |  
Db 121 NPFVNDYTWTFPSKVASKYKTANYLDHFNELHAGDSGTFGGYPDI CHDKSWDQYWLWAS 180  
| | | | |  
QY 181 QESYAAALRSIGIDAWRFDYVKGYAPWVVKDNLNMGWGWAVGEYWDTNVDVAVLNWAYSSG 240  
| | | | |  
Db 181 QESYAAALRSIGIDAWRFDYVKGYAPWVVKDNLNMGWGWAVGEYWDTNVDVAVLNWAYSSG 240  
| | | | |  
QY 241 AKVFDFALYYKMDFAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
| | | | |

Db 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGOPTIFRYDYEEWLNKDKLNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 Db 301 YAFILTYEGOPTIFRYDYEEWLNKDKLNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 QY 361 DKPGLITYINLGSSKAGRWYVYVFKFAGACIHEYTGNLGWDVKKYVSSGWVYLEAPAYDP 420  
 Db 361 SKPGLITYINLGSSKAGRWYVYVFKFAGACIHEYTGNLGWDVKKYVSSGWVYLEAPAYDP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 Db 421 ANGQYGSVWSYCGVG 436

## RESULT 6

US-10-081-872-6  
 ; Sequence 6, Application US/10081872  
 ; Publication No. US20030125534A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT APPLICATION NUMBER: US/10/081,872  
 ; PRIOR FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: Fast-Seq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 436  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated polypeptide  
 US-10-081-872-6

Query Match 97.8%; Score 2406; DB 14; Length 436;  
 Best Local Similarity 97.5%; Pred. No. 1.5e-214;  
 Matches 425; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYNDVPSGGIWDITIRKIPEDWDAGISAIWIPPSKMGKGAYS 60  
 Db 1 MAKYSELEKGGVIMQAFYNDVPSGGIWDITIRKIPEDWDAGISAIWIPPSKMGKGAYS 60  
 QY 61 MGVDYDFDLGEYDOKGTGVTETFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 Db 61 MGVDYDFDLGEYDOKGTGVTETFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNELHAGSGTGGYPPDICHDKSDQYWLWAS 180  
 Db 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNELHAGSGTGGYPPDICHDKSDQYWLWAS 180  
 QY 181 QESYAAVYLSIGIDANFDFYVKGYAPWVVKDMLNMGWAGVEYDNTVDVNLWAYSSG 240  
 Db 181 QESYAAVYLSIGIDANFDFYVKGYAPWVVKDMLNMGWAGVEYDNTVDVNLWAYSSG 240  
 QY 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTDIIWNKYPA 300  
 Db 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTDIIWNKYLA 300

QY 301 YAFILTYEGOPTIFRYDYEEWLNKDKLNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 Db 301 YAFILTYEGOPTIFRYDYEEWLNKDKLNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 QY 361 DKPGLITYINLGSSKAGRWYVYVFKFAGACIHEYTGNLGWDVKKYVSSGWVYLEAPAYDP 420  
 Db 361 DKPGLITYINLGSSKAGRWYVYVFKFAGACIHEYTGNLGWDVKKYVSSGWVYLEAPAYDP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 Db 421 ANGQYGSVWSYCGVG 436

## RESULT 7

US-10-385-305-6  
 ; Sequence 6, Application US/10385305  
 ; Publication No. US20040018607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT APPLICATION NUMBER: US/10/385,305  
 ; PRIOR FILING DATE: 2003-03-06  
 ; PRIOR APPLICATION NUMBER: US/10/081,872  
 ; PRIOR FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: Fast-Seq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 436  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated polypeptide  
 US-10-385-305-6

Query Match 97.8%; Score 2406; DB 15; Length 436;  
 Best Local Similarity 97.5%; Pred. No. 1.5e-214;  
 Matches 425; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYNDVPSGGIWDITIRKIPEDWDAGISAIWIPPSKMGKGAYS 60  
 Db 1 MAKYSELEKGGVIMQAFYNDVPSGGIWDITIRKIPEDWDAGISAIWIPPSKMGKGAYS 60  
 QY 61 MGVDYDFDLGEYDOKGTGVTETFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 Db 61 MGVDYDFDLGEYDOKGTGVTETFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNELHAGSGTGGYPPDICHDKSDQYWLWAS 180  
 Db 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNELHAGSGTGGYPPDICHDKSDQYWLWAS 180  
 QY 181 QESYAAVYLSIGIDANFDFYVKGYAPWVVKDMLNMGWAGVEYDNTVDVNLWAYSSG 240  
 Db 181 QESYAAVYLSIGIDANFDFYVKGYAPWVVKDMLNMGWAGVEYDNTVDVNLWAYSSG 240  
 QY 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTDIIWNKYPA 300  
 Db 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTDIIWNKYLA 300



Db 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHDHLAGGSTDIVVYDNDELIFVRNGY 360  
QY 361 DKPLITYINLGSSKAGRWVYVPKFGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
Db 361 SKPLITYINLGSSKAGRWVYVPKFGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
QY 421 ANGQYGSVMSYCGVG 436  
Db 421 ANGQYGSVMSYCGVG 436

RESULT 10  
US-10-081-872-38  
; Sequence 38, Application US/10081872  
; Publication No. US20030125534A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE REFERENCE: 09010-108001  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 436  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated polypeptide  
US-10-081-872-38

Query Match 96.7%; Score 2377; DB 14; Length 436;  
Best Local Similarity 96.8%; Pred. No. 7.4e-212;  
Matches 422; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MAKYSELEKGGVIMQAFYVDVPSGGIWDITROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
Db 1 MAKYLEEBEGVIMQAFYVDVPSGGIWDITROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
QY 61 MGYDPDYPFDLGEYDQGTGTETRFSGKQELVNNINTAHAYMKVIADIVINHRAGGDLEW 120  
Db 61 MGYDPDYPFDLGEYDQGTGTETRFSGKQELVNNINTAHAYMKVIADIVINHRAGGDLEW 120  
QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNHLHAGSGTFFGGYDICHDKSWDQYWLWAS 180  
Db 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNHLHAGSGTFFGGYDICHDKSWDQYWLWAS 180  
QY 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
Db 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
QY 241 AKVPDFALYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
Db 241 AKVPDFALYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
QY 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHDHLAGGSTDIVVYDNDELIFVRNGY 360  
Db 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHDHLAGGSTDIVVYDNDELIFVRNGY 360

QY 361 DKPLITYINLGSSKAGRWVYVPKFGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
Db 361 SKPLITYINLGSSKAGRWVYVPKFGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
QY 421 ANGQYGSVMSYCGVG 436  
Db 421 ANGQYGSVMSYCGVG 436

RESULT 11  
US-10-385-305-38  
; Sequence 38, Application US/10385305  
; Publication No. US20040018607A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE REFERENCE: 09010-108001  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 436  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated polypeptide  
US-10-385-305-38

Query Match 96.7%; Score 2377; DB 15; Length 436;  
Best Local Similarity 96.8%; Pred. No. 7.4e-212;  
Matches 422; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MAKYSELEKGGVIMQAFYVDVPSGGIWDITROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
Db 1 MAKYLEEBEGVIMQAFYVDVPSGGIWDITROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
QY 61 MGYDPDYPFDLGEYDQGTGTETRFSGKQELVNNINTAHAYMKVIADIVINHRAGGDLEW 120  
Db 61 MGYDPDYPFDLGEYDQGTGTETRFSGKQELVNNINTAHAYMKVIADIVINHRAGGDLEW 120  
QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNHLHAGSGTFFGGYDICHDKSWDQYWLWAS 180  
Db 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNHLHAGSGTFFGGYDICHDKSWDQYWLWAS 180  
QY 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
Db 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
QY 241 AKVPDFALYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
Db 241 AKVPDFALYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
QY 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHDHLAGGSTDIVVYDNDELIFVRNGY 360  
Db 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHDHLAGGSTDIVVYDNDELIFVRNGY 360



```
Db 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEHTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
QY 421 ANGOYGYSVMSYCGVG 436
Db 421 ANGOYGYSVMSYCGVG 436

RESULT 15
US-10-385-305-4
; Sequence 4, Application US/10385305
; Publication No. US20040018607A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kervuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated polypeptide
US-10-081-872-4

Query Match 96.3%; Score 2368; DB 14; Length 436;
Best Local Similarity 96.1%; Pred. No. 5e-211;
Matches 419; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITIRQKIPEDYDAGISAIWIPPASKMGMGAYS 60
Db 1 MAKYLEEGGLIMQAFYWDVPMGGIWDITIAQKIPDASAGISAIWIPPASKMGMGAYS 60
QY 61 MGVDPPDFDLGSDYDOKGTETRFSGKQELVNNINTAHAYGMKVIADIIVINHRAGDLEW 120
Db 61 MGVDPPDFDLGSDYDOKGTETRFSGKQELVNNINTAHAYGKVIADIIVINHRAGDLEW 120
QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
Db 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
QY 181 QESYAAALRSIGIDAWRFDYKGYAPWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
Db 181 QESYAAALRSIGIDAWRFDYKGYAPWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
QY 241 AKVDFDALYVKMDEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDI IWKYPA 300
Db 241 AKVDFDALYVKMDEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDI IWKYPA 300
QY 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360
Db 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360
QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEHTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
Db 361 SKPGLITYINLASSEAGRWVYVPKFAACIHEHTGNLGGWVDKWDSSGWVYLEAPADP 420
```

```
QY 421 ANGOYGYSVMSYCGVG 436
Db 421 ANGYIGYSVMSYCGVG 436

RESULT 15
US-10-385-305-4
; Sequence 4, Application US/10385305
; Publication No. US20040018607A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kervuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated polypeptide
US-10-385-305-4

Query Match 96.3%; Score 2368; DB 15; Length 436;
Best Local Similarity 96.1%; Pred. No. 5e-211;
Matches 419; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITIRQKIPEDYDAGISAIWIPPASKMGMGAYS 60
Db 1 MAKYLEEGGLIMQAFYWDVPMGGIWDITIAQKIPDASAGISAIWIPPASKMGMGAYS 60
QY 61 MGVDPPDFDLGSDYDOKGTETRFSGKQELVNNINTAHAYGMKVIADIIVINHRAGDLEW 120
Db 61 MGVDPPDFDLGSDYDOKGTETRFSGKQELVNNINTAHAYGKVIADIIVINHRAGDLEW 120
QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
Db 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180
QY 181 QESYAAALRSIGIDAWRFDYKGYAPWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
Db 181 QESYAAALRSIGIDAWRFDYKGYAPWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
QY 241 AKVDFDALYVKMDEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDI IWKYPA 300
Db 241 AKVDFDALYVKMDEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDI IWKYPA 300
QY 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360
Db 301 YAFILTYEGOPTIFRYDYEELNKKLKNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360
QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEHTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
Db 361 SKPGLITYINLASSEAGRWVYVPKFAACIHEHTGNLGGWVDKWDSSGWVYLEAPADP 420
```



QY 421 ANGOYGYSVWSYCGVG 436  
Db 421 ANGYTYGYSVWSYCGVG 436

Search completed: June 29, 2004, 09:41:42  
Job time : 50 secs

